Three Essays on Quantitative Asset Pricing

Dissertation
for the Faculty of Economics, Business Administration and Information Technology of the University of Zurich

to achieve the title of
Doctor of Philosophy
in Banking & Finance

presented by
Simon Scheuring
from Germany

approved in September 2012 at the request of
Prof. Dr. Felix Kübler
Prof. Dr. Karl Schmedders

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorises the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zurich, 19.09.2012

Chairman of the Doctoral Committee: Prof. Dr. Dieter Pfaff

Acknowledgements

Completing this PhD thesis was certainly the most challenging task, life has so far set upon me. The greatest and worst moments of this four-year long journey have been shared by many people.

First, and foremost, my debt of gratitude goes to my supervisor, Felix Kübler. Not only did he offer me this great opportunity, he also encouraged and supported me throughout the entire time. He freely shared his extensive knowledge and always had an open ear for all my not-yet grown research ideas. For this I cannot thank him enough.

Great thanks also go to the second member of my committee, Karl Schmedders. I have hardly ever met anyone so friendly, inspiring, knowledgeable and yet so down-to earth at the same time. I am also enormously grateful to our secretary, Ruth Häfliger, the good soul of our chair for constantly lifting our spirits. It would not have been the same without her.

Many great people crossed my path at the Swiss Banking Institute — unfortunately or rather fortunately way too many to mention. My apologies for mentioning only a selected few as representatives for all of you. Thanks Jochen for always helping out with my computer science issues; thanks Alex for relentlessly dragging me to kickboxing classes and thanks Kerstin for being my psychological and love-life consultant.

Special mentioning deserves my co-author on two papers, Benjamin. As co-authors, colleagues, office mates and friends, we have shared countless activities and arguably had the most intense relationship possible. Thanks for staying with me in good times as well as in bad times. I have learned so much from you and I am not sure whether I could have finished this thesis without you.

Finally, I would like to thank the people dearest to me: my father and mother for always believing in me, my best friends, David, Sascha, Stefan for always listening and last but not least Sylvia for entering and enriching my life.

Table of Contents

List of Figures

List of Tables

Introduction

Introduction

The objective of this dissertation is to further our understanding of the interplay between financial markets and the real economy. Since their early days financial economists have thought to understand the drivers of interest rates, asset prices and exchange rates. In a free market economy, the answer — at a general level — is extremely simple: prices are driven by the decisions of market participants, by the interplay of supply and demand. However, financial markets are arguably one of the most complex dynamic systems known to mankind. In today's globalized world, financial markets aggregate the decisions of millions of individuals and institutions. Naturally, the goal of understanding financial markets has attracted countless genius minds during the last centuries.

There is one crucial advantage we enjoy today over all those that came before us. Incredible advances in computer science and technology have provided today's researchers with entirely new possibilities. Our ancestors would not even have dared to dream of the computing power that even handheld devices provide nowadays. This thesis uses modern numerical algorithms and large scale computing power to solve and analyze complex financial economic general equilibrium models.

In spite of the large number of financial innovations introduced over time, firms and households still face many uninsurable risks. Individuals cannot fully insure themselves against unemployment or against the financial consequences of a prolonged sickness. Similarly, governments or societies typically cannot issue securities, where payoffs are contingent on the event of a recession. Thus, markets are still — and will always be — far from complete. Nevertheless, most of the asset pricing literature assumes complete markets for tractability. In complete markets, investors are able to trade on any possible future event, and hence all risks are insurable. This unrealistic assumption is probably one of the main reasons for the literature's difficulties to provide reasonable explanations for observed asset prices and the level of risk premia. In a world in which everything may be insured, people have few reasons to behave cautiously.

This observation was my main motivation for studying dynamic models with heterogeneous agents and incomplete markets. In the settings I consider, heterogeneity among agents is modeled by assuming that they face different individual stochastic income streams. Market incompleteness prevents agents from fully sharing their risks among each other. Such models are typically highly intractable and cannot be solved in closed form. Luckily, they can be rewritten as stochastic control problems for which mathematicians, computer scientists and engineers have developed powerful algorithms, iteratively combining non-linear equation solving and approximation methods. Using large-scale grid computing facilities allows to increase the scope of the models and to bring them to the data. The first two papers in this dissertation are precisely in this spirit. By applying modern techniques, I show that market incompleteness can explain some aspects of observed empirical observations of asset prices and exchange rates. The third paper complements the other two by focusing on one specific aspect of solving equilibrium models numerically, namely the approximation of exogenous processes by discrete Markov chains. The main results of the three papers can be briefly summarized as follows.

The first paper, "International Diversification and the Forward Premium", studies the relationship between interest rates, exchange rates and the real economy. Classical theory predicts that high interest rate currencies should depreciate against low interest rate currencies to account for the difference in interest earned. However, the opposite has been observed empirically. Many actors on financial markets, in particular hedge funds, have attempted to capitalize on this empirical observation by employing so-called carry trades. The idea is to borrow in a country with low interest rates, to invest in a country with high interest rates and to hope that the high interest rate currency depreciates by less than the interest rate differential. Surprisingly, this simple strategy has been extremely profitable. The paper presents a general equilibrium model that rationalizes this empirical anomaly We use heterogeneous agents to represent each country and assume that financial markets are incomplete at the international level, so that countries cannot fully insure each other against potential recessions. In addition, reflecting the psychological observation that humans get very quickly accustomed to a lifestyle, we assume that households form habits. Accordingly, drops in consumption below the level they are used to are extremely painful. Specifically, we assume that households' utility depends primarily on changes in consumption rather than on the level of consumption. We show that the combination of habit preferences and incomplete markets creates capital flows that can account for the empirical relationship between exchange rates and interest rates for ten different country pairs.

The second paper, "Asset Pricing with Idiosyncratic Risk: The Impact of Job Loss", relaxes the representative agent assumption made in most asset pricing models. The aggregation of an entire economy into a single decision maker does not reflect the situation of individual households. Even the deepest recession is no larger than a few percentage points. However, unemployment often causes households' income to drop by half or even more. Relaxing the complete markets assumption allows investigating the consequences of this imbalance between individual and aggregate risk for equilibrium asset prices. As in the first paper, market incompleteness is the crucial assumption, since it prevents agents from insuring each other and aggregating back to the representative agent. Idiosyncratic risk is captured using unemployment, which is modelled as large, long-lasting drops in households' income that arise mainly during recessions. The paper shows that unemployment risk induces sizable increases in risk premia compared to the complete markets case.

The third paper, "Multivariate Markov Chain Approximations", is a methodological paper that focuses on the relationship between the computational solution of equilibrium models and empirical data observations. Any dynamic stochastic model is driven by exogenous processes. To solve equilibrium models numerically, it is necessary to discretize these exogenous processes into discrete Markov chains. The paper presents and compares three solution approaches in the bivariate case. The first is based on ideas from numerical integration (quadrature). The second approach focuses on the moments of the process. The third approach attempts to estimate Markov chains directly from the data. These three approaches are then compared in the context of a benchmark asset pricing model. This comparison reveals that, when models involve only few discrete states, the usual choice in the literature, quadrature, is far from optimal. Furthermore, a simple experiment with moment matching demonstrates that model results can be manipulated substantially, despite matching all first and second moments precisely. Therefore, I personally recommend bin estimation as the method of choice for applied researchers. Independently, of the method, researchers have to be extremely careful and clearly document their choices when applying Markov chain approximations.

Part I

International Diversification and the Forward Premium

Time-Varying International Diversification and the Forward Premium

Benjamin Jonen, Simon Scheuring*

May 23, 2012

Abstract

This paper reproduces the slope of the uncovered interest rate parity (UIP) regression for ten country pairs within one standard deviation under rational expectations. We propose an infinite horizon dynamic stochastic general equilibrium model with incomplete markets. Heterogeneous investors experience varying risk aversion as a result of habit formation.

The underlying mechanism of the model relies on varying international diversification in the investors' portfolio choice decision. In response to their changing habit levels, investors' hedging desire varies over time. This leads to adjustments in interest rates. The habit-induced investment decisions are negatively correlated with movements in the exchange rate. This results in a negative correlation between interest rates and expected exchange rates, as implied by a negative UIP slope.

Depending on the magnitude of habits, the model is capable of reproducing positive as well as negative UIP slopes, as seen empirically in the data.

*Department of Banking and Finance, University of Zurich, benjamin.jonen@bf.uzh.ch, simon.scheuring@bf.uzh.ch. We are grateful to Felix Kubler and Karl Schmedders for continuous support. For helpful comments we would like to thank Philippe Bacchetta, Ravi Bansal, Hal Cole, Henrik Hasseltoft, Kenneth Judd, Michael Magill and Martine Quinzii as well as seminar participants at the 2010 Study Center Gerzensee Ninth Swiss Doctoral Workshop in Finance; Campus for Finance, WHU; International Conference on Operations Research, ETH; Zurich Workshop on Economics, Lucerne; and 2012 Midwest Finance Association Annual Meeting. For support in parallelizing the underlying code, we would like to thank Sergio Maffioletti and Riccardo Murri. In addition, we are thankful to the Swiss National Grid Initiative for providing computational resources.

I.1 Introduction

A large body of empirical literature[1] finds that high interest currencies tend to appreciate. This is surprising, since it implies that investors in high yield currencies benefit twice, once from the interest rate spread and once from the expected appreciation. Standard economic models predict exactly the opposite, namely that the uncovered interest rate parity (UIP) holds: high interest currencies should depreciate. The empirical phenomenon, usually referred to as the forward premium anomaly, is one of the most prevalent puzzles in international finance and has also given rise to the great popularity of carry trades[2].

Given the complexity and resilience of the puzzle[3], financial economists have been searching for a potential explanation ever since its discovery. Approaches toward a theoretical explanation emerge from three major directions: irrational expectations, market frictions or rational risk premia. This paper develops a two-country model under rational expectations without market frictions, attributing the forward premium to time-varying risk premia.

We assume that consumers form habits according to their consumption history. This changes the price of risk over time. When consumption drops close to the habit level, marginal utility increases and implied risk aversion rises. Contrarily, a large wedge between consumption and habit implies small risk aversion. Without habit, expected exchange rate (FX) appreciations always translate into a falling interest differential (confirm UIP). The introduction of habit induces shifts in investors' international diversification: Investors purchase foreign assets to hedge their consumption risk. The desire to hedge varies with different levels of income. Therefore, interest rate differentials carry time-varying risk premia for consumption growth. These risk premia are negatively correlated with FX returns. Thus, for sufficiently high habit levels, expected exchange rate appreciations can lead to increasing interest rate differentials (contradict UIP), as seen in the data.

The model's exchange rate is the ratio between tradable good prices in the two countries. We therefore assume Purchasing Power Parity holds for the tradable part of agents' income. This allows the model to generate realistic levels of inflation and FX returns simultaneously.

Markets are assumed to be incomplete on the international level. There is no asset that directly enables the representative investors to insure their income risk. This assumption is necessary to prevent countries from completely aggregating their individual risk, i.e. consume a constant percentage of the global income in tradable goods. The emerging country-specific consumption uncertainty impacts risk premia: they become larger and more varying, across time as well as across countries.

[1]The discovery is attributed to Hansen and Hodrick (1980) and Fama (1984). For surveys see Hodrick (1987) and Engel (1996).

[2]Carry trade refers to the strategy of borrowing in low interest currencies while investing in high interest currencies.

[3]For a survey see Engel (1996). Important theoretical contributions include: Alvarez, Atkeson, and Kehoe (2009), Bacchetta and van Wincoop (2010), Bansal and Shaliastovich (2009), Bekaert (1996), Colacito (2006), Farhi and Gabaix (2008) and Verdelhan (2010).

With habit levels common in the literature[4], we are able to reproduce the forward premium anomaly for ten different country pairs, composed of the five countries Australia, Germany, Japan, United Kingdom and United States. For eight out of those ten countries, the match is almost perfect, and for the remaining two the model remains within one standard deviation of the empirical observation.

This paper is related to the work of Verdelhan (2010). Verdelhan provides an explanation to the forward premium in the Campbell and Cochrane (1999) habit framework. He combines pro-cyclical interest rates with habit driven counter-cyclical risk aversion to replicate the anomaly. One restriction of his approach is that consumption has to be exogenous. In an international model, this implies the absence of trade, which Verdelhan achieves by assuming sufficiently large transportation costs. In the appendix, Verdelhan takes a first step toward a more diversified model, by reducing transportation cost and solving the planer's problem for the two countries. This paper takes the next step, by abandoning the planer and solving for a competitive equilibrium.

Thus, similarly to Verdelhan, we attribute the forward premium to rational risk premia, which vary over time due to habit formation. In our model, however, consumption is endogenous. We therefore allow for trade and international investment decisions. Theoretically, this allows for feedback effects between the two countries and generally for richer dynamics within the model. Empirically, it allows to replicate more and different moments. Most notably, we account for the low correlation between consumption and FX returns, commonly referred to as the Backus and Smith (1993) puzzle. Backus observes a disconnect between consumption and real exchange rates. Since asset prices crucially depend on correlation, matching this low correlation makes it very challenging to generate large and fluctuating risk premia. To our knowledge, this is the first model to account for the Backus and Smith (1993) puzzle (although in its nominal version), while matching negative UIP slope coefficients in a rational expectation framework.

The rest of the paper is organized as follows. In section two, we present our model, followed by a description of our numerical solution method in section three. In section four we describe our calibration. Section five discusses our model results and section six concludes.

I.2 The model

I.2.1 General setup

Real economy

This model describes an exchange economy of two infinitely lived countries, in which each country is endowed with two types of nondurable consumption goods, one tradable, one nontradable. Each country is represented by one agent[5]. In each period, agents receive a

[4]Campbell and Cochrane (1999), Verdelhan (2010).

[5]We name them agent H and agent F and they reside in country 1 and country 2.

share ϕ of their endowments in the tradable good $y_{NG,t} = \phi y_t$ and a share $1-\phi$ in the nontradable good $y_{TG,t} = (1-\phi)y_t$. The agents consider the foreign tradable good as a perfect substitute for the domestic tradable good and possess Cobb-Douglas preferences over the two consumption goods,

$$u(c_{TG,t}, c_{NG,t}) = \frac{2}{1-\gamma}\left(c_{TG,t}^{\psi} c_{NG,t}^{1-\psi}\right)^{1-\gamma},$$

where γ refers to the risk aversion and ψ to the preference for tradables. For ease of notation we refer to the vector of consumption $c_t = (c_{TG,t}, c_{NG,t})$, whenever an explicit distinction between tradables and non-tradables is not necessary.

Financial economy

Each country has separate exogenous price levels, determining the relative value of the currency. We measure the price level in terms of the nominal price of the tradable good in each country.[6] In addition, we assume that goods and assets can only be traded in the currency of the home country. Furthermore, we assume that Purchasing Power Parity (PPP) holds for tradable goods, thus determining the nominal exchange rate as

$$S_t = \frac{p_{1,t}}{p_{2,t}},$$

where $p_{1,t}$ and $p_{2,t}$ refer to the price levels (e.g. prices of tradables) in the two countries.[7]

Each country issues a one-period bond with no possibility to default. Denoting prices and nominal holdings of bonds, issued by country i by $q_{i,t}$ and $B_{i,t}$ respectively, and introducing a superscript to identify the country that chooses the economic variable, the home country's nominal budget constraint can be written as

$$C_{TG,t}^{H} \leq W_t^H + Y_{TG,t}^{H} - q_{1,t}B_{1,t}^{H} - q_{2,t}B_{2,t}^{H},$$

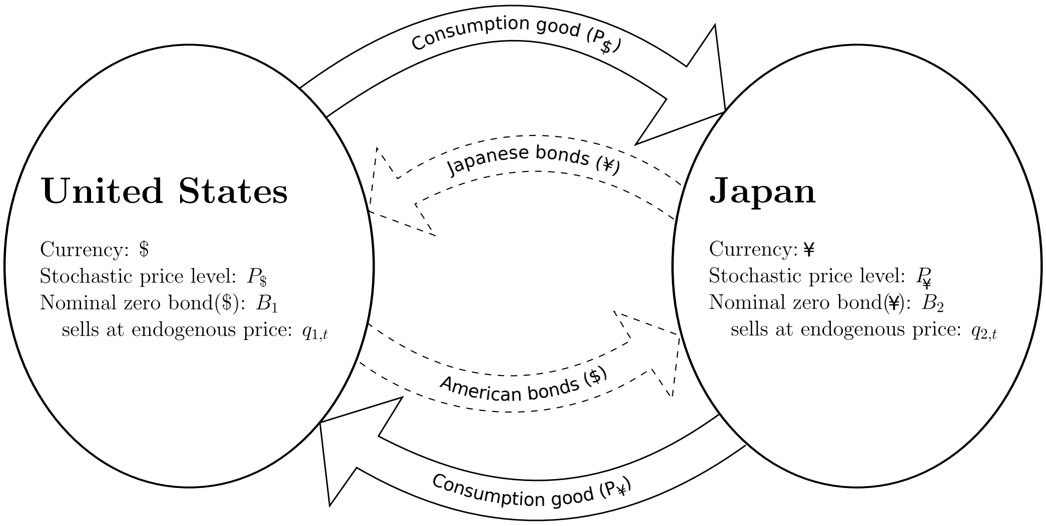
where $W_t^H = B_{1,t-1}^{H} + S_t B_{2,t-1}^{H}$ represents nominal wealth of country H. For this wealth, we assume a constant boundary on real debt

$$\frac{q_{1,t}B_{1,t} + S_t q_{2,t} B_{2,t}}{p_{1,t}} \geq \bar{w}^H.$$

[6] Prices of non-tradables have no impact on agents' decisions in our model.

[7] Given the empirical evidence on absolute and relative PPP it cannot be claimed that PPP holds for a general basket of goods. Burstein, Eichenbaum, and Rebelo (2005), however, show that PPP holds approximately for tradables, if one chooses the definition of tradable good appropriately. In particular, they distinguish between production and distribution of tradable goods. They argue that distribution is essentially nontradable. Based on this distinction they show empirically that even in times of extreme exchange rate fluctuations PPP holds approximately for tradables. Note that our model results are insensitive in the share of tradables in total income.

Figure I.1: Illustration of the market setup



This figure illustrates our model setup. Two countries trade consumption goods for nominal zero bonds, which are affected by stochastic price levels.

Uncertainty

Uncertainty enters the model through real and monetary shocks, where z_t denotes the vector of all such shocks. Shocks follow a first order Markov process with transition function $\Pi(z_{t+1}|z_t)$. Real shocks change the endowment of consumption good $(y_{NG,t}(z_t), y_{TG,t}(z_t))$ available to each country, whereas monetary shocks change the inflation rate in each country. This has two important implications. Firstly, monetary shocks determine the exchange rate through PPP. Secondly, although countries cannot default on their bonds, stochastic inflation implies a real consumption risk of holding bonds.

Note that the financial economy consists of only two bonds. Therefore, markets are generically incomplete.

Summary

Figure I.1 summarizes our model setup. Consider two countries, for example the United States and Japan. Each country has a stochastic income in its own good, a distinct currency and issues a nominally riskless zero bond. The United States sell some of their goods to Japan, while the Japanese issue bonds as a promise to repay in the future and vice versa. In equilibrium the net financial transactions will always equal the net real transactions.

Risk enters the model on the real side through stochastic income and on the financial side through stochastic inflation rates in each country, affecting the real payouts of the nominally secure bonds.

I.2.2 Habit utility

We assume investors value consumption only beyond their current habit level. The utility function, now supplemented by an external habit level[8], can be written as

$$u(c_t, h_t) = u(c_t - h_t) = u(c_{TG,t} - h_{TG,t}, c_{NG,t} - h_{NG,t}).$$

Following Constantinides (1990), Ferson and Constantinides (1991) and Heaton (1995) we specify investors' habit process[9] as a weighted average of past consumption, recursively written as

$$h_{t+1} = \rho h_t + \eta c_t. \tag{I.1}$$

For simplicity we consider the same habit level for tradables and nontradable goods, where we take nontradable consumption as a proxy for aggregate consumption in each country.[10]

This specification of habit increases the local curvature of the utility function, and thus, increases the risk aversion of agents. Moreover, risk aversion changes as agents experience different shocks to endowment. In times of consumption levels close to habit levels, marginal utilities are large and agents very risk averse. Contrarily, in times, when consumption is much higher than habit, marginal utilities are relatively small and the price of risk is low. Thus, habit formation allows for large, time varying risk premia.

Instead of calibrating the habit process parameters directly, we focus on the first two unconditional moments of the habit process, $\mathbb{E}[h]$ and $\mathbb{V}[h]$). They are more intuitive than the parameters of the habit process (I.1). The original parameters are then given by

$$\rho = \frac{\frac{\mathbb{E}[h]^2}{\mathbb{E}[c]^2}\mathbb{V}[c] - \mathbb{V}[h]}{\frac{\mathbb{E}[h]^2}{\mathbb{E}[c]^2}\mathbb{V}[c] + \mathbb{V}[h]}, \tag{I.2}$$

$$\eta = (1-\rho)\frac{\mathbb{E}[h]}{\mathbb{E}[c]},$$

where $\mathbb{E}$ refers to the unconditional expectation and $\mathbb{V}$ to the unconditional variance.

[8] The literature distinguishes internal from external habit. We follow Abel (1990) in the use of external habit formation, commonly referred to as "catching up with the Joneses".

[9] Campbell and Cochrane (1999) and Verdelhan (2010) use a nonlinear, reverse engineered habit process, which has interesting implications in their framework. However, in our opinion, Constantinides's modelling of habit is the economically more intuitive choice.

[10] The good-specific habit levels are then formed as fractions of the aggregate habit level proportionally to the amount of tradables and nontradables in the economy.

I.2.3 Optimization problem

The optimization problem for each agent is

$$\underset{C_t, B_{1,t}, B_{2,t}}{max} \quad \sum_{t=0}^{\infty} \delta^t u(c_t, h_t), \tag{I.3}$$

subject to the budget constraint, the law of motion of wealth and the borrowing constraint.

We seek a *competitive equilibrium*, that is a sequence of asset prices $q_t = (q_{1,t}, q_{2,t})$ and portfolio holdings $B_t = (B_{1,t}, B_{2,t})$[11], such that given q_t, the choice of B_t solves (I.3), subject to the agents' individual constraints and market clearing.

For each agent we can rewrite the sequence problem into the corresponding recursive problem. Define $z_t = (\pi_t^H, \pi_t^F, Y_t^H, Y_t^F)$, $\Psi_t = (W_t, z_t, h_t)$, then

$$V_t(\Psi_t) = \underset{C_t, B_{1,t}, B_{2,t}}{max} u(c_t - h_t) + \delta\, \mathbb{E}_t[V_{t+1}(\Psi_{t+1})], \tag{I.4}$$

subject to

$$\begin{aligned} C_{NG,t} &\leq Y_{NG,t}, \\ C_{TG,t} &\leq W_t + Y_{TG,t} - q_{1,t}B_{1,t} - q_{2,t}B_{2,t}, \\ W_{t+1} &= B_{1,t} + S_{t+1}B_{2,t}, \\ p_{1,t}\bar{w}^H &\leq q_{1,t}B_{1,t} + S_t q_{2,t}B_{2,t} \end{aligned}$$

In addition, we impose the following market clearing conditions:

Bonds are in zero net supply

$$\begin{aligned} B_{1,t}^H + B_{1,t}^F &= 0, \\ B_{2,t}^H + B_{2,t}^F &= 0. \end{aligned}$$

Nontradable goods cannot be traded

$$\begin{aligned} c_{NG,t}^H - y_{NG,t}^H &= 0, \\ c_{NG,t}^F - y_{NG,t}^F &= 0. \end{aligned}$$

Aggregate consumption in tradables is equal to aggregate endowments

$$c_{TG,t}^H + c_{TG,t}^F - y_{TG,t}^H - y_{TG,t}^F = 0.$$

[11] Through the budget constraint, portfolio holdings imply a consumption path.

I.2.4 Model dynamics

Incomplete markets

The only financial assets in the model are the two bonds. As they fall short of spanning the state space, markets are incomplete. Completing markets would require to introduce assets, which allow to directly insure income risk. Such assets usually do not exist in the real world. In addition, the possibility to directly insure income risk would have undesirable model implications. In the first period, the agents would negotiate to fully share their income streams. I.e. each agent would receive a constant share of the global income in tradables.

Avoiding this unrealistic implication has three impacts on risk premia: they grow larger, more volatile and differ more strongly across countries. All these features are helpful in explaining the forward premium anomaly quantitatively.[12]

Reproducing the negative slope coefficient

Figure I.2 displays the underlying mechanism in our model. The objective is to reproduce the empirical observation that $E[\Delta s]$ is negatively correlated with $i^H - i^F$. Starting from an innovation in the income process, two major channels link exchange rates and interest rate differentials. The first channel we call the UIP effect. It is the effect found in a standard economic model compatible with a slope coefficient of one.[13] The second channel is novel to our model and depends on the time-varying risk aversion induced by investors' habit formation. Changing hedging needs have the potential to reduce the slope coefficient and even make it negative. Depending on the correlation between income growth and the exchange rate, both channels have slightly different dynamics. In the majority of countries income growth is correlated with FX appreications, but in some countries with depreciations. Therefore both cases are relevant.

Negative correlation

The middle part of Figure I.2 displays the case when income growth is correlated with an *appreciation* of the home currency ($E[\Delta s] \downarrow$)[14]. Clearly, investors anticipate a currency gain and will therefore demand a lower interest rate on home bonds ($i^H \downarrow$). This is the first channel or UIP effect. The second channel is novel and provides the explanation for the existence of the forward premium. In addition to the direct effect of a positive income shock on expected exchange rates, a positive income shock also induces more consumption increasing habit formation and thus risk aversion.[15] This stimulates the home country's

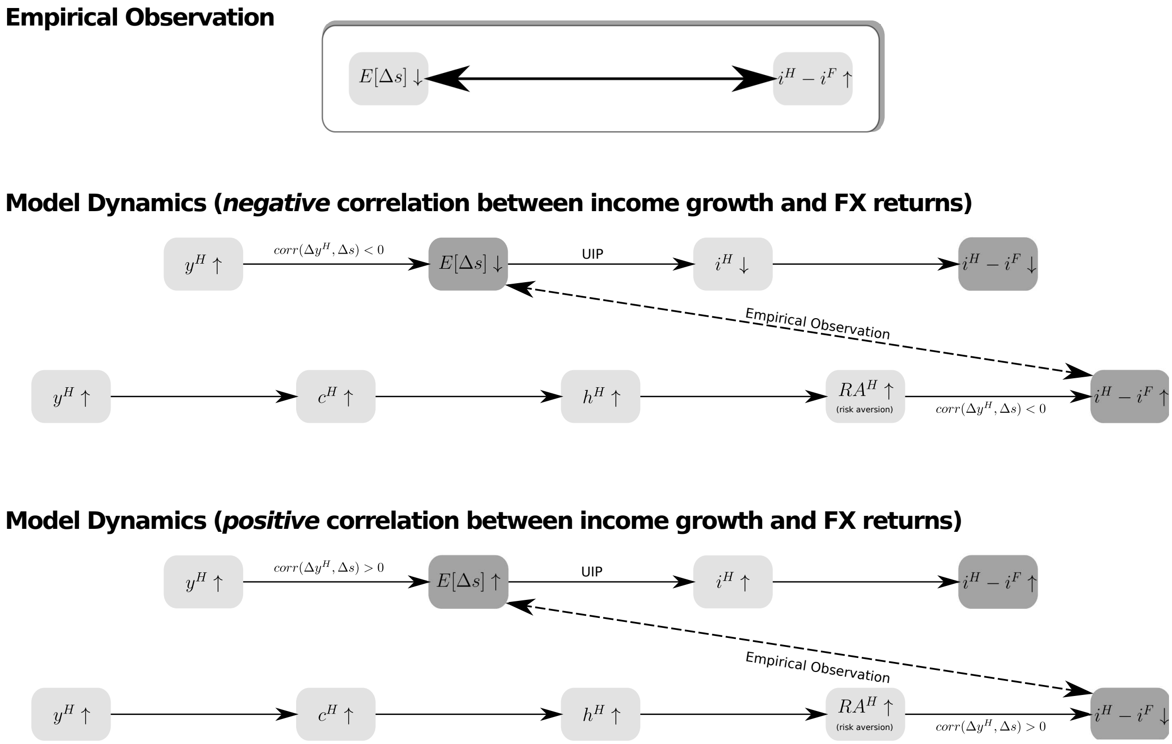
[12] See Engel (1996).

[13] See among many others: Fama and Farber (1979), Lucas (1982), Hodrick and Srivastava (1984), Hodrick and Srivastava (1986) and Engel (1992).

[14] Throughout the paper, we use the standard convention of denoting currencies as $\frac{HOME}{FOREIGN}$, therefore a decrease in the currency is equivalent to a appreciation.

[15] Risk aversion rises because habit reacts more strongly than expected consumption in our calibration. This effect is related to the question of how relative risk aversion reacts to changes in wealth. This has been

Figure I.2: Illustration of the central mechanism



This figure illustrates how the negative slope coefficient is reproduced in the model.
The upper panel states an equivalent formulation to a negative slope coefficient in the UIP regression. When the interest rate differential goes up, the expected exchange rate has to appreciate ($E[\Delta s]\downarrow$).
The middle panel shows how the puzzle can be resolved, when there is a *negative* correlation between exchange rates and income growth. The upper causality chain restates the standard UIP. The lower causality chain shows how this can be overcome by habit induced time-varying risk aversion.
The lower panel shows how the puzzle can be resolved, when there is a *positive* correlation between exchange rates and income growth. Again, the upper chain displays the UIP and the lower our habit induced risk aversion effect.

demand for foreign bonds, reduces the foreign interest rate and thus leads to a larger interest rate differential.

If the second effect quantitatively outweighs the first effect, this provides a possible explanation for the forward premium. The home currency appreciates at the same time as the interest rate differential increases.

Positive correlation
The bottom part of Figure I.2 displays the case when income growth is correlated with a *depreciation* of the home currency. Investors expect a depreciation of the home currency now. Therefore, they will demand a higher interest rate on home bonds to compensate for the expected decline in the purchasing power of their investment return. That is, according to standard theory, an FX depreciation is followed by an increasing interest differential. The mechanics of the second channel are almost identical to the negative correlation case. Higher income growth leads to habit formation and increasing risk aversion. Investors hedging desire rises. In contrast to the former case, now home assets provide a hedge against consumption risk. The interest differential falls as investors buy home bonds. Since both channels change signs, the negative slope coefficient can also be reproduced in the positive correlation case.

In summary, international diversification allows investors to hedge some of their income risk. As a result of income fluctuation and peoples' habit formation, the desire for international diversification fluctuates over time. The interest rate movements induced by this time-varying hedging desire has the potential to mitigate the UIP effect. Depending on the relative strength of both effects, the model can replicate a negative correlation between interest rate differential and expected exchange rates.

I.3 Computation

The dynamic programming problem (I.4) cannot be solved analytically. We therefore proceed to solve it computationally following methods in Judd (1998). To obtain a numerical solution the problem has to be discretized to a finite number of shocks. In practice this translates into approximating the estimated processes (i.e. income and exchange rate process for each country) by a discrete shock vector and an associated transition matrix. We simply follow the standard choice in the literature and use an implementation of Tauchen's algorithm (Tauchen (1986), Tauchen and Hussey (1991)).

We discretize the habit process into a discrete number of habit states. At the beginning of each period, habit is computed according to (I.1). If the resulting value does not lie on the grid, we replace the computed value with the habit grid's closest node.

an ongoing debate in the literature. However, recent evidence supports Arrow's original hypothesis (Arrow (1965) and Arrow (1970)) that relative risk aversion rises with higher wealth. See Halek and Eisenhauer (2001), Holt and Laury (2002) and Guiso and Paiella (2008).

Equipped with shock and transition matrices, the remaining relevant state space can be summarized by one endogenous state variable, net wealth of agent A. It summarizes the past actions of agent A. Wealth of agent B can simply be deduced through market clearing. Given the relevant state space of the economy, we use standard dynamic programming techniques to solve for the competitive equilibrium.

In particular we iterate over the agent's consumption policy. For the initial policy agents roll-over almost all of their debt, i.e. indebted agents pay back only a small amount of their loan in a two-period model. Then, in each step of the time iteration, we solve the nonlinear system of equations (see Appendix I.B.3, page 39) on a finite grid over net wealth and subsequently approximate the new consumption policy with cubic splines.

There is no theorem guaranteeing the convergence to or even the existence of a policy function satisfying the dynamic programming problem.[16] However, as long as we observe convergence toward a policy function, we know that it is a solution to the infinite horizon dynamic programming problem within the computational margin of error[17].

Finally, we simulate a large number of exogenous shocks for income and exchange rates and compute possible outcomes of the economy given the optimal policy functions. We perform the interest parity regression on the simulated data to test for the slope and observe additional implications of our model on various economic and financial variables.

I.4 Calibration

To assess our model's power to explain the forward premium anomaly, we calibrate the model to data for various countries. The set of countries, picked by historic economic significance, comprises Australia (AU), Germany (DE), Japan (JP), United Kingdom (UK) and the United States (US). The analysis puts special emphasis on the country pair United States and Japan, since these are the two largest economies, representing two dominant currencies; and most importantly as the anomaly is particularly robust for this country pair[18].

I.4.1 Data sources

For the calibration of our model, we need income growth, exchange rates, interest rates and trade shares. Except for trade shares, all data analysis is on the period from 1980 to 2010.

Income growth data is seasonally adjusted, in real terms and quarterly frequency and

[16] For a discussion see Duffie, Geanakoplos, Mas-Colell, and McLennan (1994) and Kubler and Schmedders (2005).

[17] The maximum deviations we allowed for were 10^{-10} for each individual FOC and 10^{-7} for the maximum change in consumption policies.

[18] Han (2004) performs a large cross-country, cross-period comparison to test whether the anomaly is universal. Performing regressions for varying time horizons in the range 1979 to 1998, he finds the percentage of observed negative beta coefficients to be 96% for the US and Japan.

provided by the Organisation for Economic Co-operation and Development (OECD), Eurostat and the Reserve Bank Australia.

Exchange rates are from Thomson Reuters Datastream and Eurostat, where we take the first day of each quarter in order to match quarterly income data. For the case of Germany we simply take the Euro as a proxy for Deutsche Mark.

As interest rates, we use 90 days Eurocurrency rates, again from Datastream. For Australia, there are is Eurocurrency, thus we use "Interest rate on Bank accepted bills" as provided by the Reserve Bank Australia.

Finally, we obtain trade shares for the year 1999 from the World Trade Organisation, "Share of goods and commercial services in the total trade of selected regions and economies".

I.4.2 Currency baskets

To calibrate the Markov chain, we need inflation and income data. While income data is readily available, tradable good inflation is not. Broad price indices, such as the Consumer Price Index (CPI), are not suitable since they incorporate both tradable and nontradable prices. More seriously, the usage of these indices would result in a model-implied exchange rate process that is completely different from the one observed in the data. This stands in sharp contrast to the paper's main goal of explaining the relationship between exchange rates and interest rates.

To avoid the above issues connected with price indices, we exploit the fact that the relation between tradable good prices in two countries is given as the exchange rate under PPP. More precisely, tradable good inflation in one country is measured as the valution of that country's currency against a broad index of other countrys' currencies. For each currency pair, we construct a currency basket of all remaining countries.[19] Tradable good inflation for one country is then derived as a weighted average of exchange rates of this country to all other countries in the basket. More formally, for a given country pair a, b, tradable good inflation is given as

$$\Pi_j = \sum_{\forall i \neq a,b} w_i S^{i,j} \qquad j = a, b,$$

where w_i is the weight of currency i in the basket and $S^{i,j}$ is the price of currency j in terms of currency i.

We choose the weights as the shares on world trade. More precisely, the relative value of the sum of each country's aggregate imports and exports with country i. As an example the currency basket for the country - pair US - JP is displayed in Table I.1. From here on, we will refer to the tradable good inflation process of a country simply as the country's (basket) exchange rate.

[19]It is convenient to exclude both countries in the basket in order to still obtain an exact match of the exchange rate when applying PPP.

Table I.1: **Reference currency basket — Japan - US**

Country	AU	CA	CH	DE	DK	FR	NL	NO	SE	SG	UK
Share	3.3	11.9	4.6	24.8	2.6	14.5	9.8	2.5	4.0	6.2	15.7

This table shows the composition of the reference currency basket for the country pair Japan - US. 1999 trade shares, obtained from WTO, are normalized such that they sum up to 100.

I.4.3 Estimation of exogeneous state variables

Equipped with the exchange rate process for each country, we can estimate the majority of the model parameters from data. In particular we estimate the exogenous shocks to income and inflation with a Vector Autoregressive Regression (VaR) of order one as

$$\begin{pmatrix} \Delta y_t \\ \Delta s_t \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \theta_y & \theta_{y,s} \\ \theta_{s,y} & \theta_s \end{pmatrix} \begin{pmatrix} \Delta y_{t-1} \\ \Delta s_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{t,y} \\ \epsilon_{t,s} \end{pmatrix}, \tag{I.5}$$

where Δy and Δs refer to the change in logs of income and exchange rates against the currency basket, α and θ are the estimated coefficients and ϵ residuals.

Inflation of tradables — persistence

It turns out empirically that the coefficients $\theta_{y,s}$ and θ_s are universally insignificant. Thus, none of the analyzed countries show signs of significant persistence in nominal exchange rate returns. Table I.2 displays the persistence estimates and p-values for the ten different basket currencies analysed in our model. Only one currency comes close to the significance threshold. Therefore we simply set these values for all countries to zero. Arguably, this assumption makes a difference for our model economy. Without it, a second channel for a direct payoff effect opens up, working against the above proposed habit effect. However, in our opinion, the empirical evidence legitimates the assumption of zero exchange rate persistence.

Table I.2: Persistence of FX returns

Ctry1	Ctry2	Pers.	Pval
AU	DE	-0.01	0.89
AU	JP	0.07	0.38
AU	UK	-0.05	0.52
AU	US	0.07	0.38
DE	JP	0.11	0.20
DE	UK	0.09	0.34
DE	US	0.09	0.35
JP	UK	0.13	0.11
JP	US	0.06	0.49
UK	US	0.16	0.05

Persistence estimates and p-values of log returns on exchange rates over different currency pairs.

Markov chain approximation

The remaining results of the VaR regression (I.5) need to be discretized to accomodate our model. Therefore we discretize the process for each country into a Markov chain with 9 states. Table I.3 displays various statistics describing the result of the empirical estimation for the country pair US - JP, showing the high quality of the Markov chain approximation. There are some minor deviations in standard deviations and persistences

Table I.3: Markov chain approximation — Japan - US

Ctry	Parameter			Data	[s.e.]	Model
JP	FX returns	Mean	$\mathbb{E}[\Delta s]$	1.009	0.005	1.009
		Std.	$\sigma[\Delta s]$	0.055	0.004	0.044
	Income	Mean	$\mathbb{E}[\Delta y]$	1.005	0.001	1.005
		Std.	$\sigma[\Delta y]$	0.011	0.001	0.009
		Pers.	θ_y	0.233	0.093	0.166
		Corr.	$\rho_{\Delta s,\Delta y}$	-0.208	0.089	-0.207
US	FX returns	Mean	$\mathbb{E}[\Delta s]$	0.998	0.004	0.998
		Std.	$\sigma[\Delta s]$	0.045	0.003	0.035
	Income	Mean	$\mathbb{E}[\Delta y]$	1.007	0.001	1.007
		Std.	$\sigma[\Delta y]$	0.009	0.000	0.007
		Pers.	θ_y	0.350	0.074	0.254
		Corr.	$\rho_{\Delta s,\Delta y}$	-0.030	0.079	-0.030

This table compares the first moments of the Markov chain approximation for the two exogenous process FX returns and income growth to actual data.

due to the discretization, but correlation is matched precisely. Similar accuracy is achieved for all other country pairs.

I.4.4 Remaining parameters

Some parameters, especially preference parameters cannot easily be estimated from data. Table I.4 summarizes the remaining parameters. The parameters in the top panel are picked while the parameters in the lower panel are calibrated: They are chosen as to minimize the distance between the model simulated and the empirical forward premium regression's slope coefficients.

Table I.4: Calibrated parameters

Parameter (Quarterly)		Value
Share of tradables	ϕ	0.50
Discount factor	δ	0.99
Risk aversion	γ	2.00
Preference for tradables	ψ	0.50
Average habit level	$\mathbb{E}[h]$	0.93
Habit volatitility	$\sigma[h]$	0.0057

The share of tradables in each country is set to 0.5[20], the discount factor to 0.99 on a quarterly horizon and finally the relative risk aversion to 2.00.

The habit parameterization is reported in the bottom panel. The average habit level is 0.93 and the habit volatility is 0.0057, roughly half the value of income volatility. This reflects the fact that habit is implicitly driven by changes in income, yet varies less than income.

I.5 Results

I.5.1 Simulation

Given agents' optimal policies, we simulate the model economy. The lower panel of Table I.5 displays the intercept and slope coefficient of a UIP regression using our model economy's data and corresponding actual data for the country pair US - JP. Our model matches both the slope and the intercept almost within one standard deviation. The theoretical values of the model are an approximation to the simulated value.[21]

The upper panel of Table I.5 shows the correlation between consumption growth and FX returns. It is known as the Backus and Smith (1993) puzzle (for real exchange rates), that these correlations are surprisingly low or even positive although standard economic theory would predict them to be close to -1.

Correlations are crucial in any explanation related to risk premia. Correlation directly affects covariances, which determine the stochastic discount factor and thus risk premia. Therefore, the capacity of our model to account for these low correlations is an important advantage over other risk-premia related explanations such as Bansal and Shaliastovich (2009) or Verdelhan (2010).

I.5.2 Impact of habit

Varying habit levels

Figure I.3 displays the impact of the local curvature, as implied by average risk aversion[22], on the slope coefficient for three different levels of habit volatility.[23] In the case of relatively small habit levels the model simply reproduces the uncovered interest parity, i.e. $\beta \approx 1$.

[20]Burstein, Eichenbaum, and Rebelo (2005) estimated the share of nontradables for ten different countries and found values between 0.31 and 0.57.

[21]For comparison and plotting purposes, it is inconvenient to use simulated slope coefficients, because of the introduced standard errors. Therefore, for the purpose of the regression, we make the assumption that ϵ_{t+1} is uncorrelated with time t expectations (The violation of this assumption is induced by the discretization of the state space). This allows us to compute approximate, yet exact slope coefficients, see Fama (1984).

[22]Local curvature is given by $\frac{\gamma}{c-h}$.

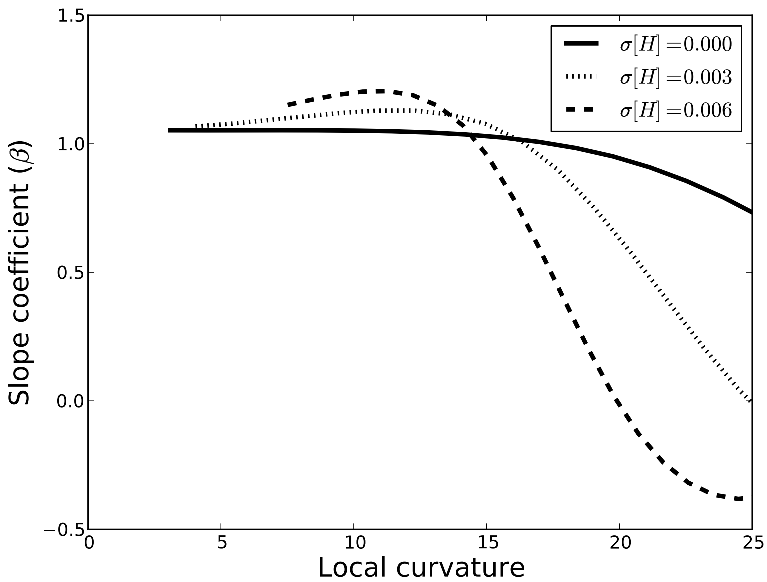
[23]The three curves have different starting values for local curvature, since combining high volatilities with low average habit levels results in negative habit persistences ρ (see (I.2) on page 17).

Table I.5: Empirical versus model implied — Japan - US

Parameter (Quarterly)	Data	s.e.	Model (sim.)	s.e.	Model (th.)
$\rho^{JP}_{\Delta s,\Delta c}$	0.22	[0.08]	0.13	[0.01]	
$\rho^{US}_{\Delta s,\Delta c}$	-0.10	[0.08]	-0.06	[0.01]	
α_{UIP}	0.03	[0.01]	0.02	[0.01]	0.02
β_{UIP}	-0.63	[0.25]	-0.36	[0.15]	-0.38

The first panel compares the model implied correlation (ρ) between real consumption growth (Δc) and FX returns (Δs) with the data. Consumption growth and FX returns are on a quarterly basis. FX returns are denoted as home over foreign, so for Japan as $\frac{¥}{\$}$ and for US as $\frac{\$}{¥}$.
The second panel compares the results of the UIP regression. α_{UIP} refers to the intercept and β_{UIP} to the slope. We report two model values. The actual simulated value with standard errors and a theoretical approximation, which we use for plotting and calibration.

Figure I.3: UIP slope coefficient — Japan - US

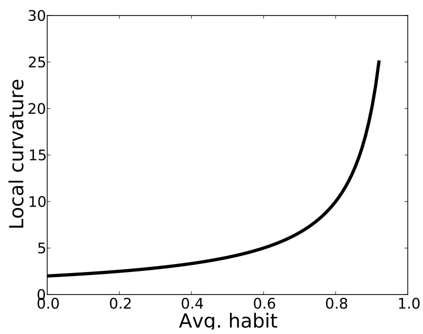


This figure displays the UIP slope for the country pair Japan - US over local curvature, as implied by average habit. The three lines represent different levels of habit volatility.

The UIP effect dominates because the habit level is too small to create large enough risk premia.

For a constant level of habit, i.e. zero habit volatility (the solid line), the UIP channel still dominates. Only at high local curvatures the habit channel starts to play a role and slowly reduces the slope coefficient. However, the coefficient remains close to one.[24] Note, that a constant habit level does not imply constant risk aversion. Since consumption varies, so does the spread $c - h$ and thus the local curvature.

Figure I.4: Local curvature



The two dotted lines in Figure 3 display cases of nonzero habit volatility. These correspond to parameterizations in which consumption shocks impact next period's habit level (i.e. $\nu \neq 0$). As average habit levels and thus risk aversion rise, the habit induced international diversification effect becomes increasingly important and finally outweighs the UIP effect. For a habit volatility of 0.003, high habit levels drive the slope coefficient down to 0. For a habit volatility of 0.006, the model predicts negative slope coefficients for average habit levels around 0.9 (i.e. an implied local curvature of 20, see Figure I.4).

I.5.3 Multiple countries

In addition to the detailed analysis for the country pair US - Japan, we apply our two-country model to nine other country pairs. These are formed by pairwise combination of Australia (AU), Germany (DE), Japan (JP), United Kingdom (UK) and the United States (US). Initially, we estimate a Markov chain for each country pair as described in the calibration section, then we solve for optimal policies and compute the model implied UIP slope coefficient. Finally, we compare these slopes to the data.

We explore three calibration scenarios. In the first scenario we pick a common habit calibration for both countries, i.e $\mathbb{E}[h_1] = \mathbb{E}[h_2]$ and $\sigma[h_1] = \sigma[h_2]$. The objective is to show that our model is in principle capable of explaining the observed forward premium for each country pair. In the second scenario, each country has its own habit parameterization, which is the same across country pairs. The idea is to infer country preferences and show that the model is able to explain the puzzle for all country pairs simultaneously. In the third calibration scenario we challenge the model with a habit process constant across all countries.

Table I.6: UIP slope coefficient — separate calibration

Ctry1	Ctry2	$\mathbb{E}[h]$	$\sigma[h]$	Model β	Emp. β	s.e.
AU	DE	0.85	0.0006	0.25	0.25	[0.21]
AU	JP	0.84	0.0074	0.12	0.10	[0.27]
AU	UK	0.81	0.0070	-0.19	-0.19	[0.23]
AU	US	0.95	0.0043	-0.04	-0.04	[0.17]
DE	JP	0.88	0.0054	0.12	0.12	[0.29]
DE	UK	0.82	0.0026	0.27	0.27	[0.18]
DE	US	0.96	0.0024	-0.04	-0.03	[0.21]
JP	UK	0.90	0.0068	-0.68	-1.05	[0.38]
JP	US	0.93	0.0057	-0.39	-0.63	[0.25]
UK	US	0.96	0.0029	-0.04	-0.04	[0.19]

This table reports the empirical vs. model implied slope coefficient for the calibration case: one habit parametrization per country pair. The first two columns refer to the country pair. The next two columns to the common habit preferences for each country pair. Finally, the last three columns compare the model implied value to the empirical observation.

Separate calibration — one habit parameterization per country pair

Table I.6 shows the result for a country pair specific habit calibration. Each country pair is analyzed separately. We assume the same habit parameterization for the two countries. The first two columns refer to the countries, the next two columns to the common habit parameterization. In the last three columns we compare the model implied slope coefficient with the empirically observed slope coefficient.

For the majority of country pairs the match is almost exact. Exceptions are Japan - US and Japan - UK. The model has difficulties reproducing these highly negative slope coefficients. Still the model implied β remains within one standard deviation for every country pair.

Simultaneous calibration — one habit parameterization per country

In this calibration, each country is assigned its own habit calibration. That is, we pick an average habit level and a habit volatility for each country to match the slope coefficients for all country pairs simultaneously. Each country has the same habit parameterization, independent of which country it is compared to. The results are reported in Table I.7. The first two columns refer to the country pairs. The next four columns display the habit parameterization for country 1 and country 2, respectively. Finally, the last three columns compare the model implied slope coefficient to the empirical observation.

Note that each country keeps its average habit and habit volatility across different country pairs. The introduced interdependencies between the different country pairs make

[24] The model becomes numerically unstable for implied local curvatures above 25. We therefore cannot report model solutions for higher levels.

Table I.7: UIP slope coefficient — simultaneous calibration

Ctry1	Ctry2	$\mathbb{E}[h_1]$	$\sigma[h_1]$	$\mathbb{E}[h_2]$	$\sigma[h_2]$	Model β	Emp. β	s.e.
AU	DE	0.86	0.0034	0.95	0.0041	-0.05	0.25	[0.21]
AU	JP	0.86	0.0034	0.88	0.0051	0.10	0.10	[0.27]
AU	UK	0.86	0.0034	0.95	0.0044	-0.07	-0.19	[0.23]
AU	US	0.86	0.0034	0.97	0.0038	0.17	-0.04	[0.17]
DE	JP	0.95	0.0041	0.88	0.0051	-0.30	0.12	[0.29]
DE	UK	0.95	0.0041	0.95	0.0044	-0.04	0.27	[0.18]
DE	US	0.95	0.0041	0.97	0.0038	-0.11	-0.03	[0.21]
JP	UK	0.88	0.0051	0.95	0.0044	-0.32	-1.05	[0.38]
JP	US	0.88	0.0051	0.97	0.0038	-0.27	-0.63	[0.25]
UK	US	0.95	0.0044	0.97	0.0038	-0.15	-0.04	[0.19]

This table reports the empirical vs. model implied slope coefficient for the calibration case: one habit parametrization per country. The first two columns refer to the country pair. The next two columns to the first two habit moments of country 1. The next two columns to the habit parameterization of country 2. Finally, the last three columns compare the model implied value to the empirical observation.

the calibration computationally much more complex.

The deviations of the slope coefficient are obviously larger than in the separate calibration. Nevertheless, every country pair remains within two standard deviations. The results from this table suggest that Americans have the highest habit level (0.97). They are closely followed by the Europeans, Germany (0.95) and United Kingdom (0.95) also display relatively high habit levels. The two countries from the far east, Japan (0.88) and Australia (0.86), show much lower habit levels.

Joint calibration — the same habit parameterization for everybody

In the final calibration exercise, we want to analyze the model's performance in the most stringent cross-country setup. We restrict the habit parameterization to be the same across all countries. The closest fit is attained for an average habit of $\mathbb{E}[h] = 0.96$ and a habit volatility of $\sigma[h] = 0.0024$. The lack of flexibility obviously results in much larger deviations of the model implied values to the actual data. While the joint calibration fails to account for JP - UK and DE - UK, the fit is acceptable for eight out of ten country pairs remaining within two standard deviations of the data.

I.6 Conclusion

This paper studies the co-movement between interest rates and exchange rates within a Lucas (1982) style model with endogenous consumption decisions. The most crucial additional assumptions are habit formation, incomplete markets and country-specific goods.

Table I.8: **UIP slope coefficient — joint calibration**

Ctry1	Ctry2	Model β	Emp. β	s.e.
AU	DE	-0.05	0.25	[0.21]
AU	JP	0.30	0.10	[0.27]
AU	UK	-0.04	-0.19	[0.23]
AU	US	0.07	-0.04	[0.17]
DE	JP	-0.20	0.12	[0.29]
DE	UK	-0.17	0.27	[0.18]
DE	US	0.37	-0.03	[0.21]
JP	UK	-0.16	-1.05	[0.38]
JP	US	-0.18	-0.63	[0.25]
UK	US	0.35	-0.04	[0.19]

This table shows the empirical vs. model implied slope coefficient for the calibration case: the same habit parameterization for every country ($\mathbb{E}[h] = 0.96$, $\sigma[h] = 0.0024$).

Theoretically, risk premia drive time-varying international hedging decisions, which lead to a possible explanation for the forward premium anomaly.

Empirically, the model convinces twofold. Firstly, it matches the first two moments for FX returns, inflation, income growth; and most notably the correlation between FX returns and income growth. Secondly, it reproduces the slope coefficient in the regression of FX returns on interest rate differentials for ten different country pairs.

Bibliography

Abel, A. B. (1990): "Asset Prices under Habit Formation and Catching up with the Joneses," *The American Economic Review*, 80(2), 38–42.

Alvarez, F., A. Atkeson, and P. J. Kehoe (2009): "Time-varying risk, Interest Rates, and Exchange Rates in General Equilibrium," *Review of Economic Studies*, 76(3), 851–878.

Arrow, K. (1965): "Aspects of the Theory of Risk Bearing, Yrjö Jahnsson Lectures," Yrjö Jahnsson Foundation, Helsinki.

——— (1970): *Essays in the Theory of Risk-Bearing*. North Holland.

Bacchetta, P., and E. van Wincoop (2010): "Infrequent Portfolio Decisions: A Solution to the Forward Discount Puzzle," *American Economic Review*, 100, 870–904.

Backus, D. K., and G. W. Smith (1993): "Consumption and real exchange rates in dynamic economies with non-traded goods," *Jounal of International Economics*, 35, 297–316.

Bansal, R., and I. Shaliastovich (2009): "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets," Working Paper.

Bekaert, G. (1996): "The Time Variation of Risk and Return in Foreign Exchange Markets: A General Equilibrium Perspective," *Review of Financial Studies*, 9(2), 427–70.

Burstein, A., M. Eichenbaum, and S. Rebelo (2005): "Large Devaluations and the Real Exchange Rate," *Journal of Political Economy*, 113, 742–784.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *The Journal of Political Economy*, 107(2), 205–251.

Colacito, R. (2006): "On the existence of the exchange rate when agents have complete home bias and non-time separable preferences," WP.

Constantinides, G. M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *The Journal of Political Economy*, 98(3), 519–543.

Duffie, D., J. Geanakoplos, A. Mas-Colell, and A. McLennan (1994): "Stationary Markov Equilibria," *Econometrica*, 62, 745–781.

Engel, C. (1992): "On the foreign exchange risk premium in a general equilibrium model," *Journal of International Economics*, 32, 305–319.

——— (1996): "The forward discount anomaly and the risk premium: A survey of recent evidence," *Journal of Empirical Finance*, 3(2), 123–192.

Fama, E. F. (1984): "Forward And Spot Exchange Rates.," *Journal of Monetary Economics*, 14(3), 319 – 338.

Fama, E. F., and A. Farber (1979): "Money, Bonds and Foreign Exchange," *The American Economic Review*, 69, 639–649.

Farhi, E., and X. Gabaix (2008): "Rare Disasters and Exchange Rates: A Theory of the forward premium puzzle," WP.

Ferson, W. E., and G. M. Constantinides (1991): "Habit persistence and durability in aggregate consumption: Empirical tests," *Journal of Financial Economics*, 29(2), 199 – 240.

Guiso, L., and M. Paiella (2008): "RISK AVERSION, WEALTH, AND BACKGROUND RISK," *Journal of the European Economic Association*, 6(6), 1109–1150.

Halek, M., and J. G. Eisenhauer (2001): "Demography of Risk Aversion," *The Journal of Risk and Insurance*, 68(1), pp. 1–24.

Han, B. (2004): "Is the forward premium puzzle universal?," *Applied Economics Letters*, 11, 131–134.

Hansen, L. P., and R. J. Hodrick (1980): "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88, 829–853.

Heaton, J. (1995): "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications," *Econometrica*, 63, 681–717.

Hodrick, R. J. (1987): *The empirical evidence on the efficiency of forward and futures foreign exchange markets*. Harwood Academic Publishers.

Hodrick, R. J., and S. Srivastava (1984): "An Investigation of Risk and Return in Forward Foreign Exchange.," *Journal of International Money and Finance*, 3, 166 – 190.

——— (1986): "The Covariation of Risk Premiums and Expected Future Spot Exchange Rates," *Journal of International Money and Finance*, 5, 5–21.

Holt, C. A., and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *The American Economic Review*, 92(5), pp. 1644–1655.

Judd, K. L. (1998): *Numerical Methods in Economics*. The MIT Press.

Kubler, F., and K. Schmedders (2005): "Approximate versus Exact Exquilibria in Dynamic Economies," *Econometrica*, 73, 1205–1235.

Lucas, R. E. (1982): "Interest Rates and Currency Prices in a Two-Country World," *Journal of Monetary Economics*, 10, 335–359.

Tauchen, G. (1986): "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20(2), 177 – 181.

Tauchen, G., and R. Hussey (1991): "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59(2), 371–396.

Verdelhan, A. (2010): "A Habit-Based Explanation of the Exchange Rate Risk Premium," *The Journal of Finance*, 1, 123–146.

Zangwill, W. I., and C. B. Garcia (1981): *Pathways to Solutions, Fixed Points and Equilibria*. Prentice-Hall.

I.A Robustness in technical parameters

Table I.9: Technical parameters

Parameter (Quarterly)		Value
Habit boundaries scale	ζ	1
Habit grid size	n_h	3
Wealth boundaries	$\bar{w}^H = \bar{w}^F$	-0.1
Wealth grid size	n_w	11

In addition to our economic calibration, there are also a few technical parameters, which we need to choose for the numerical procedure. These parameters are listed in Table I.9. The parameter choice reflects a trade-off between computational effort and accuracy. The idea of this section is to show that our main model result, the UIP slope coefficient, does not dramatically change in any of these parameters.

Habit discretization

We discretize habit in the following fashion. Adding (subtracting) habit volatility times the habit boundary scale factor ζ from average habit yields the upper (lower) bound for the habit grid. Given the boundaries, we construct a linearly spaced grid with n_h points. Since it is convenient to have average habit as a gridpoint, we restrict the number of gridpoints to an uneven number,

Figures I.5a and I.5b display the change in the slope coefficient when varying these parameters. The number of grid points has almost no impact while the scaling factor has a slight impact on the slope coefficient. Different discretization change the actual volatility of habit resulting in different slope coefficients. Since habit volatility is calibrated to fit the UIP slope, this lack of robustness is not a major issue. It only limits the comparability of the absolute level of habit volatility across different numbers of habit nodes (n_h).

Wealth discretization

Wealth is also discretized on an equally spaced grid. The boundaries are set to $\bar{w}$. n_w determines the number of grid point. Figures I.5c and I.5d clearly show that both parameters have no major impact on our model result.

I.B First order conditions

I.B.1 Normalization

To solve our model we first rewrite the nominal problem (eq. I.4) into the corresponding real problem. For this purpose we set price level of nontradables to 1 and the price level of tradables as p_1 respectively p_2 for each country.

Let us denote $R_{1,t} = \frac{1}{1+\pi_t^H}$ and $R_{2,t} = \frac{1}{1+\pi_t^F}$ as the real returns of each bond. Furthermore we redefine the shock vector and state space in real terms as follows: $z_t = (R_{1,t}, R_{2,t}, y_t^H, y_t^F)$ and $\Psi_t = (w_t, z_t, h_t)$. Then the dynamic programming problem transforms into

$$V_t(\Psi_t) = \underset{c_t, b_{1,t}, b_{2,t}}{max} u(c_t, h_t) + \delta\,\mathbb{E}_t[V_{t+1}(\Psi_{t+1})],$$

subject to

$$\begin{aligned} c_{TG,t} &\leq w_t + y_{TG,t} - q_{1,t}b_{1,t} - q_{2,t}b_{2,t}, \\ c_{NG,t} &\leq y_{NG,t}, \\ w_{t+1} &= R_{1,t+1}b_{1,t} + R_{2,t+1}b_{2,t}, \\ b_{1,t} &\geq \frac{\bar{b_1}}{E[R_{1,t+1}]}, \\ b_{2,t} &\geq \frac{\bar{b_2}}{E[R_{2,t+1}]}, \\ \bar{w} &\leq b_{1,t}q_{1,t} + b_{2,t}q_{2,t}, \\ c_{TG,t} &\geq h_{TG,t}. \end{aligned}$$

The last inequality is unnecessary in theory. The utility function is simply not defined for values smaller than 0. However, it is necessary to enforce the condition for computational reasons, as a solver might try to evaluate the function for $c_{TG,t} < h_{TG,t}$. Depending on the choice of the risk aversion, this could either result in complex numbers or even lead to a potential solution of the equation system, with no economic meaning.

To facilitate computation, we additionally normalize each agent's problem with factors κ^A and κ^B respectively. Given homothetic preferences the individual's policies are simply scaled by the normalization factor. Thus, the equilibrium remains unchanged under the appropriate adjustment of market clearing conditions (see next section).
Define $\tilde{z}_t = (R_{1,t+1}, R_{2,t+1}, \tilde{y}_t^H, \tilde{y}_t^F)$ and $\tilde{\Psi}_t = (\tilde{w}_t, \tilde{z}_t, \tilde{h}_t)$ where $\kappa\tilde{y}_t = y_t$, then

$$V_t(\tilde{\Psi}_t) = \underset{\tilde{c}_t, \tilde{b}_{1,t}, \tilde{b}_{2,t}}{max} u(\tilde{c}_t, \tilde{h}_t) + \delta\,\mathbb{E}_t[V_{t+1}(\tilde{\Psi}_{t+1})],$$

subject to

$$
\begin{aligned}
\tilde{c}_{TG,t} &\leq \tilde{w}_t + \tilde{y}_{TG,t} - q_{1,t}\tilde{b}_{1,t} - q_{2,t}\tilde{b}_{2,t}, \\
\tilde{c}_{NG,t} &\leq \tilde{y}_{NG,t}, \\
\tilde{w}_{t+1} &= R_{1,t+1}\tilde{b}_{1,t} + R_{2,t+1}\tilde{b}_{2,t}, \\
\tilde{b}_{1,t} &\geq \frac{\tilde{\bar{b}}}{E[R_{1,t+1}]}, \\
\tilde{b}_{1,t} &\geq \frac{\tilde{\bar{b}}_1}{E[R_{1,t+1}]}, \\
\tilde{b}_{2,t} &\geq \frac{\tilde{\bar{b}}_2}{E[R_{2,t+1}]}, \\
\tilde{\bar{w}} &\leq \tilde{b}_{1,t}q_{1,t} + \tilde{b}_{2,t}q_{2,t}, \\
\tilde{c}_{TG,t} &\geq \tilde{h}_{TG,t}
\end{aligned}
$$

I.B.2 Kuhn-Tucker Conditions

Concavity of the utility function allows us to impose equality for the first two conditions. Inserting conditions two and three and denoting for simplicity $u(c_t, h_t) = u(c_{TG,t})$ we can write the Lagrangian as

$$
\begin{aligned}
\mathcal{L} = u(\tilde{c}_{TG,t}) + \delta\, \mathbb{E}_t[V_{t+1}(\tilde{\Psi}_{t+1})] \;&+\; \mu\ (\tilde{y}_{TG,t} + \tilde{w}_t - q_{1,t}\tilde{b}_{1,t} - q_{2,t}\tilde{b}_{2,t} - \tilde{c}_{TG,t}) \\
&+\; \lambda_1\ (\tilde{b}_{1,t}E_t[R_{1,t+1}] - \tilde{\bar{b}}_1) \\
&+\; \lambda_2\ (\tilde{b}_{2,t}E_t[R_{2,t+1}] - \tilde{\bar{b}}_2) \\
&+\; \lambda_3\ (\tilde{b}_{1,t}q_{1,t} + \tilde{b}_{2,t}q_{2,t} - \tilde{\bar{w}}) \\
&+\; \lambda_{nn}\ (\tilde{c}_{TG,t} - \tilde{h}_{TG,t}).
\end{aligned}
$$

Deriving the Lagrangian with respect to each choice variable, adding the conditions and restrictions on the Lagrange multipliers provides us with the following system of first order conditions:

$$\frac{\partial \mathcal{L}}{\partial \tilde{c}_{TG,t}} = u'(\tilde{c}_{TG,t}) - \mu + \lambda_{nn} \overset{!}{=} 0,$$

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\tilde{b}_{1,t}} &= \delta \mathbb{E}_t \left[\frac{\partial V_{t+1}}{\partial \tilde{b}_{1,t}} \right] - \mu q_{1,t} + \lambda_1 E[R_{1,t+1}] + \lambda_3 q_{1,t} \\
&= \delta \sum_{z_{t+1} \in \Gamma(z_t)} \left[\pi(z_{t+1}|z_t) \frac{\partial u}{\partial \tilde{c}_{TG,t+1}} \frac{\partial \tilde{c}_{TG,t+1}}{\partial \tilde{w}_{t+1}} \frac{\partial \tilde{w}_{t+1}}{\partial \tilde{b}_{1,t}} \right] - \mu q_{1,t} + \lambda_1 E[R_{1,t+1}] + \lambda_3 q_{1,t} \\
&= \delta \sum_{z_{t+1} \in \Gamma(z_t)} \left[\pi(z_{t+1}|z_t) \frac{\partial u}{\partial \tilde{c}_{TG,t+1}} \frac{\partial \tilde{c}_{TG,t+1}}{\partial \tilde{w}_{t+1}} R_{1,t+1} \right] - \mu q_{1,t} + \lambda_1 E[R_{1,t+1}] + \lambda_3 q_{1,t} \overset{!}{=} 0, \\
\frac{\partial \mathcal{L}}{\tilde{b}_{2,t}} &= \delta \mathbb{E}_t \left[\frac{\partial V_{t+1}}{\partial \tilde{b}_{2,t}} \right] - \mu q_{2,t} + \lambda_2 + \lambda_3 q_{2,t} \\
&= \delta \sum_{z_{t+1} \in \Gamma(z_t)} \left[\pi(z_{t+1}|z_t) \frac{\partial u}{\partial \tilde{c}_{TG,t+1}} \frac{\partial \tilde{c}_{TG,t+1}}{\partial \tilde{w}_{t+1}} \frac{\partial \tilde{w}_{t+1}}{\partial \tilde{b}_{2,t}} \right] - \mu q_{2,t} + \lambda_2 E[R_{2,t+1}] + \lambda_3 q_{2,t} \\
&= \delta \sum_{z_{t+1} \in \Gamma(z_t)} \left[\pi(z_{t+1}|z_t) \frac{\partial u}{\partial \tilde{c}_{TG,t+1}} \frac{\partial \tilde{c}_{TG,t+1}}{\partial \tilde{w}_{t+1}} R_{2,t+1} \right] - \mu q_{2,t} + \lambda_2 E[R_{2,t+1}] + \lambda_3 q_{2,t} \overset{!}{=} 0,
\end{aligned}$$

$$\begin{aligned}
\lambda_1(\tilde{b}_{1,t} E[R_{1,t+1}] - \tilde{\bar{b}}_1) &= 0, \\
\lambda_2(\tilde{b}_{2,t} E_t[R_{2,t+1}] - \tilde{\bar{b}}_2) &= 0, \\
\lambda_3(\tilde{b}_{1,t} q_{1,t} + \tilde{b}_{2,t} q_{2,t} - \bar{w}) &= 0 \\
\lambda_{nn}(\tilde{c}_{TG,t} - (1-\eta) h_t) &= 0,
\end{aligned}$$

$$\lambda_1 \geq 0, \quad \lambda_2 \geq 0, \quad \lambda_3 \geq 0, \quad \lambda_{nn} \geq 0$$

where $\Gamma(z_t)$ denotes all states possibly following z_t and $\pi(z_{t+1}|z_t)$ are the transition probabilities.

The same set of equations exists for the second agent and is completed by the market clearing conditions

$$\begin{aligned}
\kappa^A \tilde{b}^A_{1,t} + \kappa^B \tilde{b}^B_{1,t} &= 0, \\
\kappa^A \tilde{b}^A_{2,t} + \kappa^B \tilde{b}^B_{2,t} &= 0, \\
\kappa^A \tilde{c}^A_{TG,t} + \kappa^B \tilde{c}^B_{TG,t} &= \kappa^A \tilde{y}^A_{TG,t} + \kappa^B \tilde{y}^B_{TG,t}.
\end{aligned}$$

The market clearing conditions apply to the unnormalized economy. Thus, terms are unnormalized with the agent specific normalization coefficient.

I.B.3 Alternative Conditions

It is computationally inconvenient to work with the inequality constraints for the Lagrange multiplier. Therefore we use the following reformulation as described in Zangwill and Garcia (1981).

The key is to replace the Lagrange multipliers by slacks, which are decomposed into a positive and negative part

$$\alpha^+ = [\ max(0, \alpha)\ ]^k,$$
$$\alpha^- = [\ max(0, -\alpha)\ ]^k.$$

One would expect a k of 2 or 3 to work best to avoid any kinks in the nonlinear system of equations. However, surprisingly, we find that $k = 1$ outperforms any other choice.

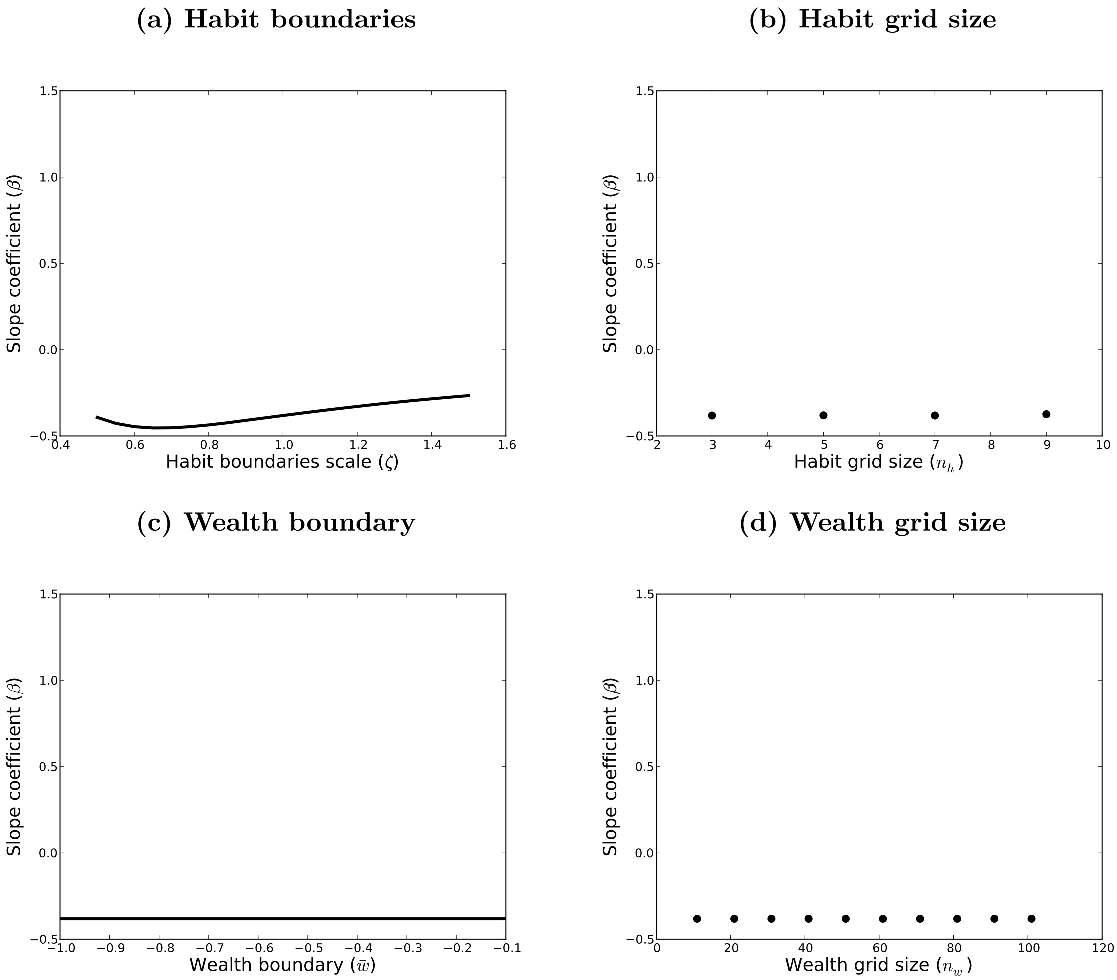
This allows us to rewrite the first order conditions into the following equivalent system

$$\frac{\partial \mathcal{L}}{\partial \tilde{c}_{TG,t}} = u'(\tilde{c}_{TG,t}) - \mu + \alpha_{nn}^+ + \qquad \stackrel{!}{=} 0,$$
$$\frac{\partial \mathcal{L}}{\tilde{b}_{1,t}} = \delta \sum_{z_{t+1} \in \Gamma(z_t)} \left[\pi(z_{t+1}|z_t) \frac{\partial u}{\partial \tilde{c}_{TG,t+1}} R_{1,t+1} \right] - \mu q_{1,t} + \alpha_1^+ E_t[R_{1,t+1}] + \alpha_3^+ q_{1,t} \quad \stackrel{!}{=} 0,$$
$$\frac{\partial \mathcal{L}}{\tilde{b}_{2,t}} = \delta \sum_{z_{t+1} \in \Gamma(z_t)} \left[\pi(z_{t+1}|z_t) \frac{\partial u}{\partial \tilde{c}_{TG,t+1}} R_{2,t+1} \right] - \mu q_{2,t} + \alpha_2^+ E_t[R_{2,t+1}] + \alpha_3^+ q_{2,t} \quad \stackrel{!}{=} 0,$$

$$\alpha_1^- - (\tilde{b}_{1,t} E[R_{1,t+1}] - \tilde{\tilde{b}}_1) = 0,$$
$$\alpha_2^- - (\tilde{b}_{2,t} E_t[R_{2,t+1}] - \tilde{\tilde{b}}_2) = 0,$$
$$\alpha_3^- - (\tilde{b}_{1,t} q_{1,t} + \tilde{b}_{2,t} q_{2,t} - \bar{w}) = 0,$$
$$\alpha_{nn}^- - (\tilde{c}_{TG,t} - (1 - \eta) h_t) = 0,$$

α can be interpreted as the shadow price of the borrowing constraint. If the constraint does not bind then α is negative and α^- positive which equalizes the $\geq$ constraint. Thus, essentially the borrowing constraint does not have a shadow price. If α is positive α^+ is positive showing up in the FOCs while the borrowing constraint exactly binds. The higher α the more costly is the constraint.

Figure I.5: UIP slope coefficient over various technical parameters



These figures report robustness checks in four technical parameters. We show how changes in these parameters affect the slope coefficient in the base calibration.

Part II

Asset Pricing with Idiosyncratic Risk: The Impact of Job Loss

Asset Pricing with Idiosyncratic Risk: The Impact of Job Loss

Benjamin Jonen, Simon Scheuring*

May 23, 2012

Abstract

This paper studies the impact of unemployment risk on risk premia in an incomplete markets economy with many infinitely-lived heterogeneous agents. Job loss is modeled as large, but rare, persistent idiosyncratic shocks with heteroskedastic countercyclical volatility. Within an otherwise standard model and despite conservative assumptions on preferences, we simultaneously generate a sizeable equity premium and a low risk-free rate.

*Department of Banking and Finance, University of Zurich, benjamin.jonen@bf.uzh.ch, simon.scheuring@bf.uzh.ch. We are highly indebted to Felix Kubler for pointing us towards the idea and continuous support throughout the project. We would also like to thank Karl Schmedders for thoughtful comments. We are very grateful to Johannes Brumm for a detailed discussion of a previous draft. For support in parallelizing the underlying code we would to like to thank Sergio Maffioletti and Riccardo Murri. We are thankful to the Swiss National Grid Initiative for providing computational resources. We gratefully acknowledge financial support from NCCR-FINRISK.

II.1 Introduction

Mehra and Prescott (1985) show that the representative-agent complete markets model cannot replicate essential empirical facts in finance. An important strand of the literature has identified idiosyncratic risk as a potential explanation for the observed asset prices[1]. However, attempts to model idiosyncratic risk, generally involve modeling heterogeneous agents with incomplete markets. Since it is usually necessary to track one state variable per agent, such models quickly become intractable. One approach in the literature has been to reduce the state space by assuming the endogenous policy functions to be independent of each other.[2] This obviously remediates the problem of intractability, however, arguably also avoids the multidimensional nature of the problem.

In this paper we follow a different strand of literature[3] and employ the approximation algorithm of Smolyak (1963). It breaks the curse of dimensionality by picking interpolation points in a clever way. As a result, computing time grows only polynomially rather than exponentially as the number of state variables increases. This allows us to analyze more involved models of idiosyncratic risk with up to six agents without simplifying assumptions on policies. In particular, we are able to model job loss. By definition, idiosyncratic risk has to cancel out on the aggregate level. Thus, the loss of one agent must be the gain of the others. If one attempts to model unemployment in a two-agent economy then the employment income will be unrealistically large, which has severe implications on asset prices. Extending the analysis to a larger number of agents mitigates this problem and makes an analysis of a skewed income distribution feasible.

Within a Lucas (1978) framework we incorporate several model extensions suggested in the literature of idiosyncratic risk. Mankiw (1986) finds that the more concentrated shocks are on a small part of the population, the higher the risk premium. This is the case for job loss, which we model as large, but rare, idiosyncratic shocks. Among many others, Lucas (1994) and Heaton and Lucas (1996) stress the importance of market frictions. We assume that markets are dynamically incomplete: No asset allows direct insurance of income shocks and when unemployed, agents face a tight borrowing constraint. The constraint prevents them from smoothing consumption through borrowing in bad times and repaying in good times. As Constantinides and Duffie (1996) we rely on persistent idiosyncratic shocks with heteroskedastic countercyclical volatility. Empirically, unemployment is a lagged indicator of economic growth. Our model takes this co-movement into account by assuming that unemployment risk is high in recessions and low in booms and that agents remain unemployed until the economy picks up again.

The novel combination of these features generates realistic risk premia, despite low aggregate income growth volatility and conservative assumptions on preferences. A realistic calibration for the United States results in an equity premium of 4.7% with a risk-free rate

[1] See among many others Bewley (1982), Mankiw (1986), Weil (1992), Telmer (1993), Lucas (1994), Heaton and Lucas (1996) or Constantinides and Duffie (1996).

[2] See Aiyagari (1994), Krusell and Smith (1997), Krusell and Smith (1998), Storesletten, Telmer, and Yaron (2007).

[3] Krueger and Kubler (2004), Krueger and Kubler (2006) and Malin, Krueger, and Kubler (2011).

of 1.4%. Thus, the model generates large risk premia despite low risk-free rates.

The paper is organized as follows: Section 2 describes the model. Section 3 specifies the unemployment dynamics. The calibration is found in section 4, followed by the results in section 5. Finally, section 6 concludes.

II.2 The model

II.2.1 Economy

We consider an endowment economy, populated by n infinitely-lived agents. We denote average income per agent as Y_t and assume aggregate income (nY_t) to grow with a stochastic rate $g_{t+1} = \frac{Y_{t+1}}{Y_t}$. The state of the economy can be summarized by the wealth vector $W_t = (W_t^1, W_t^2, \ldots, W_t^n)$ and the current shock z_t. The current shock describes the aggregate state of the economy as well as the individual level of income.

II.2.2 Preferences

Each agent has the same recursive Epstein and Zin (1989) preferences over consumption C_t

$$V_t(W_t, z_t) = \left[(C_t - \varsigma Y_t)^{1-\rho} + \beta \mathbb{E}_t \left[V_{t+1}(W_{t+1}, z_{t+1})^{1-\gamma} \right]^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}}, \tag{II.1}$$

where $\psi = \frac{1}{\rho}$ is the intertemporal elasticity of substitution (IES), γ the risk aversion, β the subjective time discount factor and ς the relative subsistence level of consumption as a fraction of average income (Y_t). The subsistence level captures the idea that households need a minimum level of consumption to survive. More concretely, we think of the subsistence level as the level of consumption necessary to provide a household with basic needs as discussed in Sharif (1986). Utility is then derived only from consumption which exceeds the basic needs of survival.

II.2.3 Assets

One firm produces the entire output nY_t of the economy. The firm liquidates at the beginning of each period and splits the output into payoffs to employees ($Y_{l,t}$), bond holders ($P_{b,t}$) and stock holders ($P_{s,t}$)

$$Y_t = Y_{l,t} + P_{b,t}\bar{b} + P_{s,t}\bar{s},$$

where $\bar{b}$ and $\bar{s}$ denote the supply of bonds and stocks.

Typically, the claims of bond holders and employees are senior to the claims of stock holders. Abstracting from the possibility of default, we follow an extension in Mehra and Prescott (1985) and define stock payoffs as the stochastic part of the economy

$$P_{s,t}\bar{s} = Y_t - (1 - \bar{s})\mathbb{E}_{t-1}[Y_t].$$

Then aggregate wages and payouts to bond holders are fractions of the anticipated output of the economy

$$P_{b,t} = \mathbb{E}_{t-1}[Y_t],$$
$$Y_{l,t} = \bar{l}\mathbb{E}_{t-1}[Y_t],$$

where $\bar{l}$ is the share of wages as part of the expected aggregate firm output.

To acquire claims on output next period, people can invest in the firm at the end of each period by buying bonds (b_t^i) or investing in stocks (s_t^i). In total all claims need to equal the amount available to distribute

$$\sum_{i=1}^{n} b_t^i = \bar{b},$$
$$\sum_{i=1}^{n} s_t^i = \bar{s}.$$

II.2.4 Idiosyncratic risk

Similar to Lucas (1994), we introduce idiosyncratic shocks to agents' income Ψ_t^i. Then, agents' labor income can be decomposed into the non-stochastic part ($Y_{l,t}$) and an idiosyncratic shock: $Y_{l,t}^i = Y_{l,t} + \Psi_t^i$. To match aggregate income, the idiosyncratic shocks need to sum up to zero $\sum_{i=1}^{n} \Psi_t^i = 0$.

Then, the budget constraint is

$$W_t^i + Y_{l,t}^i = Q_{s,t}s_t^i + Q_{b,t}b_t^i + C_t^i \tag{II.2}$$

and wealth accumulates according to

$$W_{t+1}^i = P_{s,t+1}s_t^i + P_{b,t+1}b_t^i. \tag{II.3}$$

Furthermore, we assume that agents face state-contingent borrowing constraints. End of period net wealth of agent i (investment) has to lie above some minimum fraction of average income $\underline{i}_t^i < 0$

$$Q_{s,t}s_t^i + Q_{b,t}b_t^i \geq \underline{i}_t^i Y_t.$$

II.3 Unemployment dynamics

The previous section described the general framework. In this section, we specify the structure of idiosyncratic shocks. While they are part of the model assumptions, we devote an entire section to unemployment dynamics for two reasons: First, to reflect the importance

Table II.1: Aggregate Markov chain

(a) Shocks

State	g_t
1	$\mu_g - \sigma_g$
2	$\mu_g + \sigma_g$

(b) Transition

State	1	2
1	p	$1-p$
2	$1-p$	p

idiosyncratic shocks play for the emerging model predictions. Second, to freely discuss computational considerations and impacts on calibration choices.

To make models numerically tractable, it is necessary to discretize the aggregate and individual shock space. The usual approach is to write the model in continuous states and then apply a discretization method, such as Tauchen and Hussey (1991). However, in this paper, we choose to model the dynamics of exogenous processes directly in a discrete shock space.

This leads to the usual trade-off between computational feasibility (a small number of states) and a realistic setting (a large number of states). The idea behind our Markov chain is to minimize the number of states, while maintaining the necessary components to model aggregate and idiosyncratic risk.

II.3.1 Markov chain

Similar to Mankiw (1986), our Markov chain is composed of two parts: Aggregate shocks indicate the distribution of average income over time. Idiosyncratic shocks specify the distribution of income across agents.

Aggregate Markov chain

Table II.1 shows the discretization of the aggregate economy's growth rate (g_t) into a Markov chain with two states. Economic growth is high in one state and low in the other. To match the first two unconditional moments, we construct the shock matrix by subtracting and adding the observed standard deviation to the observed mean. The corresponding transition matrix is parameterized by p, denoting the probability to remain in the same growth state. Empirically, income growth persistence is small (see Table II.5). Thus, for simplicity, we assume i.i.d income growth, i.e. $p = 0.5$.

Individual Markov chain

We model job loss through a separate Markov chain displayed in Table II.2. The shock matrix specifies idiosyncratic shocks and consists of n rows, where each row i represents the state in which agent i falls unemployed. Each agent's shock to income is represented by one column in the matrix. The entries display the percent deviation of agent i's labor income from average income, denoted $\psi_t^i = \frac{\Psi_t^i}{Y_{l,t}}$. Agent i suffers from job loss in state i and

Table II.2: **Individual Markov chain**

State	ψ_t^1	ψ_t^2	ψ_t^3	$\dots$	ψ_t^n
1	$-\Delta$	$\frac{\Delta}{n-1}$	$\frac{\Delta}{n-1}$	$\dots$	$\frac{\Delta}{n-1}$
2	$\frac{\Delta}{n-1}$	$-\Delta$	$\frac{\Delta}{n-1}$	$\dots$	$\frac{\Delta}{n-1}$
3	$\frac{\Delta}{n-1}$	$\frac{\Delta}{n-1}$	$-\Delta$	$\dots$	$\frac{\Delta}{n-1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	$\frac{\Delta}{n-1}$	$\frac{\Delta}{n-1}$	$\frac{\Delta}{n-1}$	$\dots$	$-\Delta$

thus receives a lower income. Agent j receives a small positive income adjustment in state i to cancel out agent i's shock at the aggregate. For example, in state 1 agent 1 receives an income shock of $\psi_t^1 = -\Delta$. Thus agent 1's labor income amounts to $Y_{l,t}^1 = (1-\Delta)Y_{l,t}$. Similarly agent 2's income in state 1 is increased by $\frac{\Delta}{n-1}$ resulting in a labor income of $Y_{l,t}^2 = \left(1 + \frac{\Delta}{n-1}\right) Y_{l,t}$.

Common Markov chain

Finally, Table II.3 combines the individual and the aggregate into a common Markov chain. First, unemployment only occurs in an economic downturn. Second, once unemployed, an agent remains unemployed with probability p and regains employment when the economy recovers with probability $1-p$. Since job loss only occurs in an economic downturn, idiosyncratic risk is countercyclical and heteroskedastic.

This construction has two advantages: Economically, it incorporates countercyclical heteroscedastic idiosyncratic shocks. Computationally, the number of states is the number of agents plus one, thus as small as possible.

II.3.2 Intuition

The next two sections show that with the above structure of unemployment dynamics, the model is capable of simultaneously generating large risk premia and low risk-free rates with modest and conservative calibration choices. In this subsection we attempt to provide an intuition for this result.

Unemployment constitutes a large, potentially long lasting, hit on agents' income. From a partial equilibrium perspective, agents generally have three ways to smooth consumption as a response to idiosyncratic shocks. First, agents can buy direct ex-ante insurance. Second, agents can save as a precaution. Third, agents can indebt themselves, whenever income is low and pay it back later.

In our model, agents cannot buy ex-ante insurance against future job losses, since markets are dynamically incomplete. Indebting is limited through the borrowing constraint. Therefore, agents have to rely mainly on precautionary savings to smooth their consumption. Figure II.1 illustrates the development of wealth over the unemployment cycle. Initially, the agent will build up wealth as a precaution for a potential future job

Table II.3: Common Markov chain

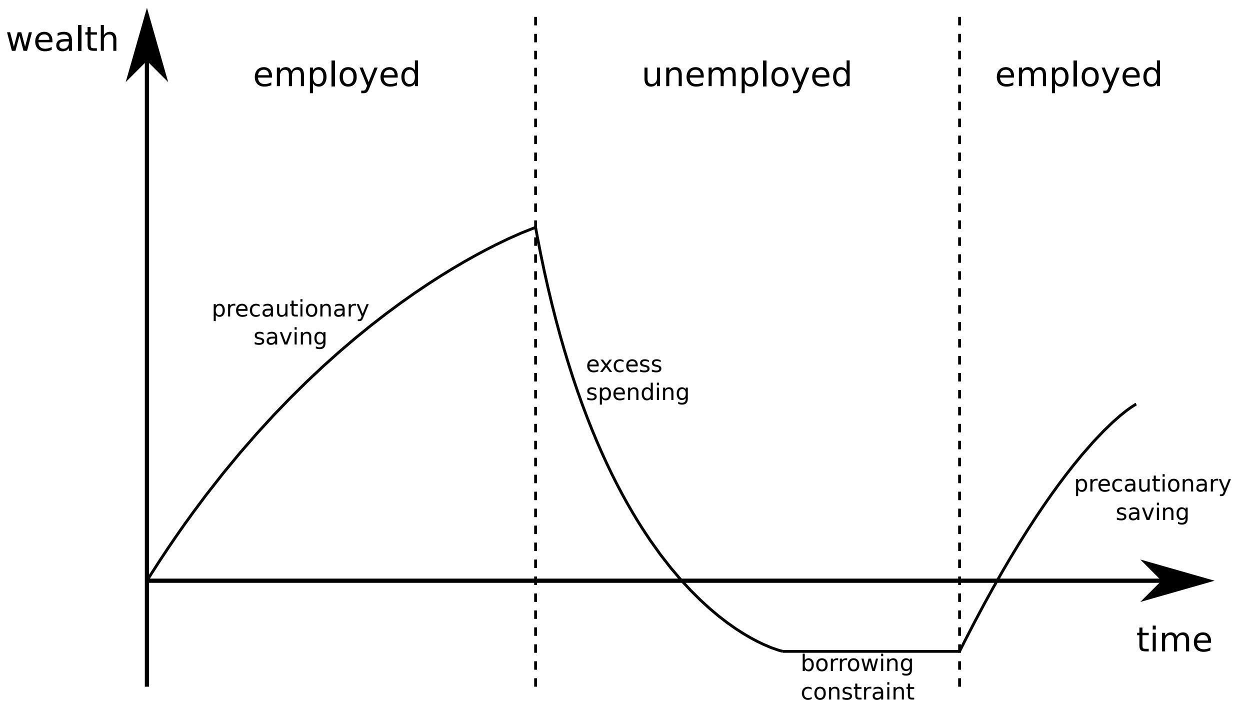
(a) Shocks

State	g_t	ψ_t^1	ψ_t^2	ψ_t^3	$\dots$	ψ_t^n
1	$\mu_g - \sigma_g$	$-\Delta$	$\frac{\Delta}{n-1}$	$\frac{\Delta}{n-1}$	$\dots$	$\frac{\Delta}{n-1}$
2	$\mu_g - \sigma_g$	$\frac{\Delta}{n-1}$	$-\Delta$	$\frac{\Delta}{n-1}$	$\dots$	$\frac{\Delta}{n-1}$
3	$\mu_g - \sigma_g$	$\frac{\Delta}{n-1}$	$\frac{\Delta}{n-1}$	$-\Delta$	$\dots$	$\frac{\Delta}{n-1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	$\mu_g - \sigma_g$	$\frac{\Delta}{n-1}$	$\frac{\Delta}{n-1}$	$\frac{\Delta}{n-1}$	$\dots$	$-\Delta$
n+1	$\mu_g + \sigma_g$	0	0	0	$\dots$	0

(b) Transition

State	1	2	3	$\dots$	n	$n+1$
1	p	0	0	$\dots$	0	$1-p$
2	0	p	0	$\dots$	0	$1-p$
3	0	0	p	$\dots$	0	$1-p$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
n	0	0	0	$\dots$	p	$1-p$
$n+1$	$\frac{1-p}{n}$	$\frac{1-p}{n}$	$\frac{1-p}{n}$	$\dots$	$\frac{1-p}{n}$	p

Figure II.1: Illustration of wealth dynamics



loss. Upon job loss, the agent enters a phase of excess spending, consuming the accumulated wealth. After all savings are depleted, agents start borrowing. Eventually, however, the borrowing constraint will hit. From this point on, consumption is limited to income. When the economy picks up again and the agent regains employment, the cycle starts over from the beginning with precautionary saving.

All in all, precautionary savings is the primary way for agents to smooth idiosyncratic income shocks. High asset demand raises prices and reduces returns, allowing us to generate large risk-premia while maintaining a realistically low risk-free rate.

II.4 Calibration

This section discusses the model calibration. Table II.4 summarizes discretionary choices, while the top panel of Table II.5 shows estimated input data.

II.4.1 Population size

The economy is populated by six agents ($n = 6$). According to the specification of the unemployment dynamics six agents imply an average unemployment rate of 8.3%. As a comparison, the US post war average is 5.8%. Figure II.2 displays the historical evolution. One can see large fluctuations over time, ranging from below 3% to almost 10%. The empirical annual volatility is 1.6%. Our model implied annual volatility is somewhat larger with 4.2%.

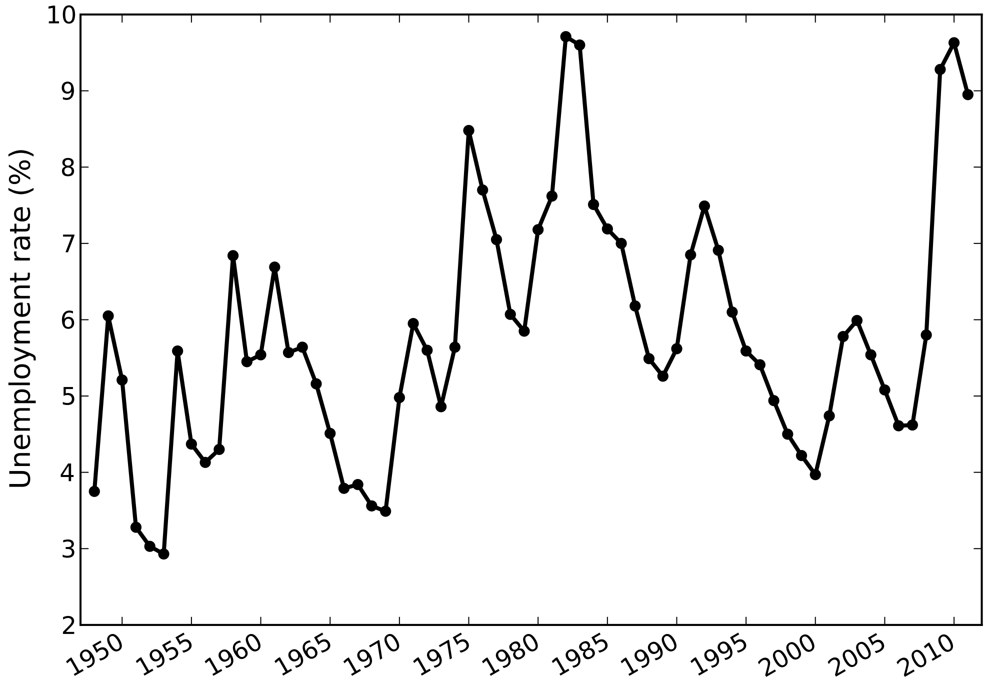
II.4.2 Preferences

We distinguish two calibration choices in the second panel of Table II.4. The third column specifies the parameterization under constant relative risk aversion (CRRA), that is when $\psi = \frac{1}{\gamma}$. We set risk aversion (γ) to 5. Then we adjust the discount factor (β) to arrive at reasonable values for the risk-free rate. In particular we set $\beta = 0.97$.

Since the risk premium under this specification is still relatively small, we investigate the effect of increasing γ (reducing ψ). Increasing risk aversion reduces ψ in the CRRA framework. Empirically ψ is estimated to be "significantly different from zero, and probably close to 1"[4]. Thus, as we increase risk aversion it seems plausible to move away from the standard CRRA preferences and adjust ψ. The fourth column describes the parameterization under the general Epstein-Zin preferences where we set risk aversion to 8 and the intertemporal elasticity of substitution to 0.33. Again, to arrive at a reasonable risk-free rate we adjust the discount factor, in this case 0.99, to arrive at a reasonable value for the risk-free rate.

[4]Beaudry and Wincoop (1996). Hansen and Singleton (1982) and Campbell and Cochrane (1999) also find the IES to be larger than $\frac{1}{\gamma}$.

Figure II.2: Unemployment rate



Seasonally adjusted unemployment rate over time, 16 years and older, annual averages, provided by the U.S. Department of Labor.

Table II.4: Parameters

Parameter		CRRA	EZ
Number of agents	n	6	6
Discount factor	β	0.97	0.99
Risk aversion	γ	5	8
IES	ψ	0.20	0.33
Subsistence level	ς	10%	10%
Bond supply	$\bar{b}$	20%	20%
Stock supply	$\bar{s}$	15%	15%
Borrowing constraint, when unem.	$\underline{i}$	-5%	-5%
Replacement rate	$1-\Delta$	45%	45%

This table displays two calibration choices. The left column represents preferences with constant relative risk aversion (CRRA), $\gamma = 1/\psi$. The right column represents Epstein-Zin (EZ) preferences, $\gamma \neq 1/\psi$.

We set the subsistence level (ς) in both parameterizations to 10%. Thus, we assume that the US median-income household of $\$49,400$[5] cannot survive with less than $\$5,000$.

II.4.3 Financial economy

The third panel of Table II.4 displays the parameterization of the financial economy. The firm in our economy distributes its production to the workforce, bond and stock holders. We assume these shares to be 65% for labor ($\bar{l}$), 20% for bonds ($\bar{b}$) and 15% for stocks ($\bar{s}$). We also assume agents face a borrowing constraint, denoted $\underline{i}$. We enforce this constraint only in the unemployment state. This is computationally easier and does not matter economically since agents in employment have no reason to borrow. In the real world unemployed agents have a hard time borrowing beyond negative net wealth. In this situation credit cards tend to be one of the few ways to borrow but without a proper proof of employment limits tend to be tight. Heaton and Lucas (1996) argue a value between 0% and -10% is reasonable. We pick the middle and choose the boundary on net wealth of the unemployed ($\underline{i}$) as -5%.

II.4.4 Real economy

Aggregate growth

The first two panels of Table II.5 depict aggregate income and consumption growth. Income growth refers to the real gross domestic product, seasonally adjusted from the U.S. Bureau of Economic Analysis over the horizon 1947 - 2010. Consumption is private final consumption expenditure, also seasonally adjusted over the same time horizon, obtained from the Organisation for Economic Co-operation and Development (OECD). Due to the lack of a savings technology, aggregate income is identical to aggregate consumption. Matching income growth as opposed to consumption growth is an arbitrary choice.

Unemployment

Calibrating unemployment requires choosing two parameters: first, the *replacement rate* $(1 - \Delta)$, how much income unemployed agents receive relative to how much they received in employment; second, the *duration of unemployment*, the average time it takes agents to regain employment.

Figure II.3a shows the replacement rate time series for the United States from 1988 to 2010. In 1988 the replacement rate was about 44% increasing to about 47% in the last decade. In a longer perspective, government transfers have increased over time. Therefore, historically the replacement rate has certainly been much lower. We take the conservative value of 45% as the replacement rate $(1 - \Delta)$. Thus, falling unemployed implies an income drop of 55%.

[5] DeNavas-Walt, Proctor, and Smith (2011).

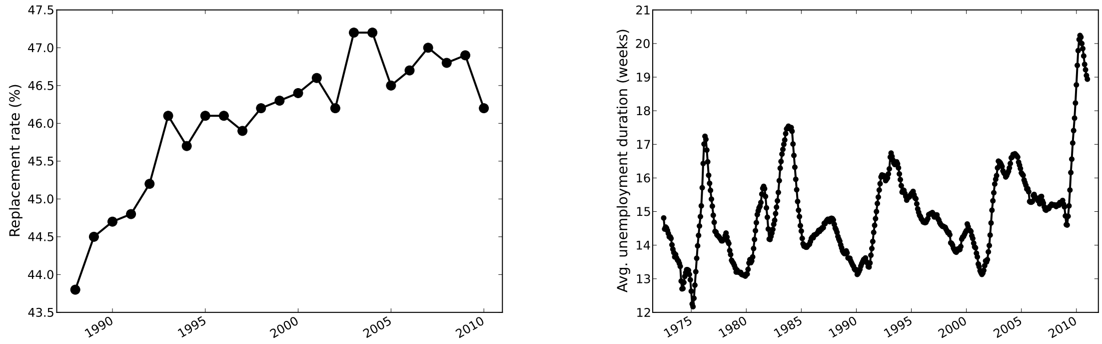
Table II.5: Moments — base calibration

Parameter		Data	CRRA	EZ
Avg. income growth	μ_g	3.3%	3.3%	3.3%
Income growth volatility	σ_g	2.8%	2.8%	2.8%
Income growth AC	$AC[g]$	−0.1%	0%	0%
Avg. consumption growth	μ_g	3.4%	3.3%	3.3%
Consumption growth vola.	σ_g	2.0%	2.8%	2.8%
Consumption growth AC	$AC[g]$	9.9%	0%	0%
Idiosyncratic income vola	$\sigma\left[Y^i_{t+1}/Y^i_t\right]$	25.1%	40.0%	40.0%
Idiosyncratic cons. vola	$\sigma\left[C^i_{t+1}/C^i_t\right]$	6% − 12%	14.4%	15.2%
Avg. market return	$\mathbb{E}[R^m]$	8.7%	5.1%	6.1%
Market return vola.	$\sigma[R^m]$	17.1%	18.9%	19.1%
Avg. risk-free rate	$\mathbb{E}[R^f]$	1.4%	1.6%	1.4%
Risk-free rate vola.	$\sigma[R^f]$	2.5%	2.4%	1.7%
Risk premium	$\mathbb{E}[R^m - R^f]$	7.3%	3.5%	4.7%

This table compares model implied moments with empirical observations. The two calibrations for CRRA and EZ preferences are displayed in the two rightmost columns. The data sources are as follows:
Income growth: BEA, Real Gross Domestic Product, seasonally adjusted, 1947-2010.
Consumption growth: OECD, private final consumption expenditure, 1947-2010.
Idiosyncratic income volatility: From Heaton and Lucas (1996).
Idiosyncratic consumption volatility: From Brav, Constantinides, and Geczy (2002).
Market return: Value weighted NYSE, including dividends, 1947-2010, deflated by CPI from CRSP.
Risk-free rate: T-Bills 90 days, deflated by CPI from CRSP.

Figure II.3: Unemployment

(a) Replacement rate **(b) Average benefit collection time**



Replacement rate: "Average Weekly UI [unemployment insurance] Benefit as a Percent of Average Weekly Wage", US Department of Labor.
Average benefit collection time: "The average number of weeks for which unemployment insurance claimants collect benefits under regular state programs", US Department of Labor.

Autocorrelation in US aggregate income growth is very small (see Table II.5). Thus, we model income growth as a random walk. As the model is computed in quarterly frequency, this implies, by construction of the Markov chain, an expected unemployment time of 6 months[6]. It is unclear which empirical proxy for average unemployment time to look at. One possibility is to look at average benefit collection time, see Figure II.3b. Historically, benefit collection time has been between 13 and 20 weeks. However, as many people in unemployment cannot find a job until government transfers run out, the average duration of unemployment is certain to be higher. Thus, we believe, the implied value of 6 months, e.g. about 25 weeks is reasonable.

II.5 Results

In this section we first discuss the results under the main calibration. Next we discuss how the model results are affected by changes in the calibration and attempt to demonstrate the underlying mechanisms. For this purpose, we present three scenarios. First, we analyze the implications of Epstein-Zin preferences by varying risk aversion and IES. Second, we display comparative statics by presenting changes in a single input parameter. Third, we look at the impact of job loss by considering the alternative scenario of no idiosyncratic risk.

[6] Every 3 months an unemployed agent has a chance $p = 0.5$ to regain employment. Thus, the expected unemployment time is $\sum_{i=1}^{\infty} p^i 3i = 3\frac{p}{(1-p)^2} = 6$ months.

Table II.6: Asset prices with Epstein-Zin

γ	ψ	β	$\mathbb{E}[R^m]$	$\mathbb{E}[R^f]$	$\sigma[R^m]$	$\sigma[R^f]$	$\mathbb{E}[R^m - R^f]$
5	0.20	0.97	5.1%	1.6%	18.8%	2.4%	3.5%
5	0.33	0.97	9.9%	6.6%	20.1%	1.2%	3.3%
8	0.20	0.97	−1.9%	−6.5%	17.1%	3.2%	4.6%
8	0.33	0.97	7.2%	2.4%	19.1%	1.9%	4.8%
5	0.20	0.97	5.1%	1.6%	18.8%	2.4%	3.5%
5	0.33	1.04	4.3%	1.1%	19.2%	0.9%	3.1%
8	0.20	0.86	4.7%	−0.2%	17.8%	4.3%	4.8%
8	0.33	0.99	6.1%	1.4%	19.1%	1.7%	4.7%
	Data		8.7%	1.4%	17.1%	2.5%	7.3%

This table shows asset pricing moments for varying combinations of risk aversion (γ), intertemporal elasticity of substitution (ψ) and the discount factor (β). The first panel keeps the discount rate constant. The second panel varies the discount rate to obtain reasonable values for the risk-free rate. The last panel repeats the empirical observation from Table II.5.

II.5.1 Base calibration

Table II.5 displays empirical moments obtained from simulating the model under the two base calibrations for CRRA and EZ discussed in the previous chapter. The second panel displays aggregate consumption. As we do not model a production side, aggregate consumption is identical to aggregate income. Therefore, while income is matched, consumption moments obviously deviate from the empirical ones.

The third panel shows idiosyncratic income and consumption volatility. The respective empirical moments are taken from the literature (Heaton and Lucas (1996) and Brav, Constantinides, and Geczy (2002)). While slightly larger, the model implied moments are reasonably close.

The fourth panel displays the main result of the paper: Averages and volatilities of market return and risk-free rate and the implied risk premium. In both cases the risk-free rate and volatilities are very close to the empirical values. For the CRRA case the model generates a risk premium of 3.5%, with the Epstein-Zin calibration the risk premium is 4.7%.

II.5.2 Epstein-Zin implications

We would like to understand in more detail the effects of Epstein-Zin preferences. For this purpose, Table II.6 displays different combinations of risk aversion and intertemporal elasticity of substitution. We distinguish two cases. In the upper panel we keep the discount rate constant at $\beta = 0.97$. In the lower panel, we adjust the discount rate to keep an almost constant risk-free rate.

Let us first look at the upper panel with a constant discount rate. An increase in ψ

Table II.7: Preferences — sensitivities

Case	$\mathbb{E}[R^m]$	$\mathbb{E}[R^f]$	$\sigma[R^m]$	$\sigma[R^f]$	$\mathbb{E}[R^m - R^f]$
Base	5.1%	1.6%	18.8%	2.4%	3.5%
β = 0.96 (0.97)	6.0%	2.5%	19.0%	2.4%	3.5%
β = 0.98 (0.97)	4.2%	0.8%	18.7%	2.3%	3.4%
γ = 4 (5)	9.5%	6.6%	19.9%	1.6%	2.9%
γ = 6 (5)	−0.7%	−4.6%	17.6%	3.0%	3.8%
ψ = 0.15 (0.2)	1.2%	−2.3%	18.2%	2.8%	3.5%
ψ = 0.25 (0.2)	7.8%	4.4%	19.4%	1.9%	3.4%
ς = 0 (10%)	11.3%	7.8%	20.2%	1.9%	3.5%
$\bar{s}$ = 10% (15%)	5.5%	0.4%	28.7%	3.2%	5.0%
$\bar{b}$ = 10% (20%)	1.6%	−1.8%	17.2%	3.7%	3.4%

The first line repeats the model results for the CRRA case from Table II.5. The following lines show deviations in one parameter. The first column describes the parameters, in parentheses we repeat the respective value in the base case. The following columns display the different moments for each calibration. β discount factor; γ risk aversion; ψ intertemporal elasticity of substitution; ς subsistence level; $\bar{s}$ stock supply; $\bar{b}$ bond supply.

implies a larger tolerance of different consumption levels across time — agents will smooth their consumption less, therefore the risk-free rate volatility falls. Furthermore, when we increase ψ, agents have less incentives for precautionary savings — demand for bonds and stock falls and interest rates as well as the market return rise. As ψ determines intertemporal choices, the impact on the risk premium is rather small. A change in γ increases volatilities and risk premium, since agents are more afraid of risk. An increased risk aversion also decreases the risk-free rate. Agents will have a greater fear of unemployment. They rely on precautionary savings to prevent losses and therefore asset prices rise, e.g. returns fall.

In the lower panel, we keep the risk-free rate almost constant to separate the effects more clearly. ψ has strong effects on the risk-free rate volatility, yet only minor indirect effects on the risk premium. An increase in risk aversion increases volatilities and the risk premium.

II.5.3 Sensitivities in preference parameters

Table II.7 shows how the model responds to changes in one of the input parameters. The first two lines repeat moments of the CRRA column of Table II.5 as a reference. In the following lines, we vary one input parameter from Table II.4 at a time. The value in parentheses repeats the respective value of the base case. We will now discuss the effects one by one.

The discount rate (β) works as expected. A lower discount rate implies a higher interest rate and vice versa. The effects on volatilities and the risk premium are negligible. The risk

Table II.8: Impact of job loss

$1-\Delta$	γ	β	$\mathbb{E}[R^m]$	$\mathbb{E}[R^f]$	$\sigma[R^m]$	$\sigma[R^f]$	$\mathbb{E}[R^m - R^f]$
100%	2	0.90	24.5%	22.9%	24.0%	1.1%	1.6%
100%	2	0.99	13.2%	11.8%	21.9%	1.0%	1.4%
100%	5	0.90	46.1%	41.6%	28.7%	1.8%	4.5%
100%	5	0.99	33.0%	28.8%	26.1%	1.6%	4.1%
45%	2	0.90	19.4%	17.8%	22.4%	0.3%	1.6%
45%	2	0.99	8.8%	7.4%	20.4%	0.3%	1.4%
45%	5	0.90	11.6%	7.9%	19.9%	2.8%	3.7%
45%	5	0.99	3.4%	0.0%	18.6%	2.3%	3.4%
	Data		8.7%	1.4%	17.1%	2.5%	7.3%

This table compares different replacement rates $(1-\Delta)$. In the upper panel, the replacement rate is 100%, e.g. job loss has no effect. In the lower panel, the replacement rate is 45%, e.g. in case of unemployment, income drops by 55%. The lines report moments of asset prices for different values of risk aversion (γ) and (β) for preferences with constant relative risk aversion. The last line repeats the empirical observations from Table II.5 as a reference.

aversion (γ) and IES (ψ) show similar qualitative results as in Table II.6. Risk aversion decreases returns and increases risk premia. The IES affects primarily the returns and has only minor effects on risk premia. The subsistence level (ς) has a strong impact on asset returns. Agents are forced to a minimum level of consumption, thus unemployment is particularly painful. As a precaution, agents invest more in assets, e.g. asset prices rise and returns fall.

Reducing the stock supply ($\bar{s}$) implies an increase in the stock volatility. In our model, all aggregate risk is carried by stock holders. As there are less stocks, the relative risk increases. Therefore, the stock volatility increases to 28.7% and consequently the risk premium increases to 5.0%. The second effect is common for stock and bond supply ($\bar{b}$). Less supply, implies less possibilities to save, e.g. less supply of assets in general. As the supply of assets decreases, the price increases, e.g. asset returns fall.

II.5.4 Impact of job loss

Table II.8 shows the impact of the introduction of job loss. In the upper panel, we show possible model calibrations for the discount rate and risk aversion without idiosyncratic risk. Since, we calibrated aggregate risk on income rather than consumption, introduced a subsistence level and leveraged our firm, the model is capable of creating sizable risk premia even without idiosyncratic risk. However, at the cost of extremely large risk-free rates. With this result, we are in line with Mehra and Prescott (1985) and Weil (1989), who show that when β is constrained to lie below 1, it is impossible to generate large risk premia, while keeping interest rates at a reasonable level. In this sense, we find the puzzle to be more about the historically low interest rate, rather than the equity premium.

The introduction of idiosyncratic risk in the lower panel opens the path for large risk premia and realistic risk-free rates. Since agents fear unemployment, they have strong incentives to save as a precaution. Therefore, asset prices rise and returns decrease to reasonable values. By calibrating $\gamma = 5$ and $\beta = 0.99$, we obtain a zero interest rate and still maintain a sizable risk premium. Therefore, our model results suggest, that the large risk premia are derived from aggregate risk, however, individual risks and the induced precautionary savings justify the historically low interest rate and answer the question raised by Weil (1989) "why is the risk-free rate so low?".

II.6 Conclusion

This paper relates risk premia to unemployment risk. Without idiosyncratic risk, we find it possible to generate large risk premia, but only at the cost of an unreasonably large interest rate. Similar to the literature[7], we thus conclude the equity premium puzzle to be more about the question why risk-free rates have been so low, rather than why equity returns have been so high. Unemployment risk provides an answer. When unemployed households face tight credit constraints and incomplete markets prevent them from insuring their idiosyncratic risk, they have to rely on precautionary savings to dampen the effects of unemployment. This creates strong demand for bonds and causes interest rates to fall to realistic levels.

[7] See Kocherlakota (1996) for a survey.

Bibliography

Aiyagari, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 109(3), pp. 659–684.

Beaudry, P., and E. v. Wincoop (1996): "The Intertemporal Elasticity of Substitution: An Exploration using a US Panel of State Data," *Economica*, 63(251), pp. 495–512.

Bellman, R. (1957): *Dynamic Programming*. Princeton University Press.

Bewley, T. F. (1982): "Thoughts on Tests of the Intertemporal Asset Pricing Model," Working Paper, Northwestern University.

Brav, A., G. M. Constantinides, and C. C. Geczy (2002): "Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence," *Journal of Political Economy*, 110, 793–824.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *The Journal of Political Economy*, 107(2), 205–251.

Constantinides, G. M., and D. Duffie (1996): "Asset Pricing with Heterogeneous Consumers," *The Journal of Political Economy*, 104, 219–240.

DeNavas-Walt, C., B. D. Proctor, and J. C. Smith (2011): "Income, Poverty, and Health Insurance Coverage in the United States: 2010," Current Population Reports — United States Census Bureau.

Epstein, L. G., and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57(4), 937–969.

Hansen, L. P., and K. J. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50(5), pp. 1269–1286.

Heaton, J., and D. J. Lucas (1996): "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *The Journal of Political Economy*, 104, 443–487.

JUDD, K. L. (1998): *Numerical Methods in Economics*. The MIT Press.

KOCHERLAKOTA, N. R. (1996): "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, 34(1), pp. 42–71.

KRUEGER, D., AND F. KUBLER (2004): "Computing equilibrium in OLG models with stochastic production," *Journal of Economic Dynamics and Control*, 28(7), 1411 – 1436.

——— (2006): "Pareto-Improving Social Security Reform when Financial Markets are Incomplete!?," *American Economic Review*, 96(3), 737–755.

KRUSELL, P., AND A. A. J. SMITH (1997): "Income and wealth heterogeneity, portfolio choice, and equilibrium asset returns," *Macroeconomic Dynamics*, 1, 387–422.

——— (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896.

LUCAS, D. J. (1994): "Asset pricing with undiversifiable income risk and short sales constraints: Deepening the equity premium puzzle," *Journal of Monetary Economics*, 34, 325–341.

LUCAS, R. E. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429–1445.

MALIN, B. A., D. KRUEGER, AND F. KUBLER (2011): "Computing Stochastic Dynamic Economic Models with a Large Number of State Variables: A Description and Application of a Smolyak-Collocation Method," *Journal of Economic Dynamics and Control*, 35, 229–239.

MANKIW, N. G. (1986): "The Equity Premium and the Concentration of Aggregate Shocks," *Journal of Financial Economics*, 17, 211–219.

MEHRA, R., AND E. C. PRESCOTT (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15(2), 145 – 161.

SHARIF, M. (1986): "The Concept and Measurement of Subsistence: A Survey of the Literature," *World Development*, 14(5), 555 – 577.

SMOLYAK, S. (1963): "Quadrature and Interpolation Formulas for Tensor Products of Certain Classes of Functions," *Soviet Mathematics, Doklady*, 4, 240–243.

STOKEY, N. L., AND R. E. LUCAS (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press.

STORESLETTEN, K., C. I. TELMER, AND A. YARON (2007): "Asset pricing with idiosyncratic risk and overlapping generations," *Review of Economic Dynamics*, 10, 519–548.

TAUCHEN, G., AND R. HUSSEY (1991): "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59(2), 371–396.

TELMER, C. I. (1993): "Asset-Pricing Puzzles and Incomplete Markets," *The Journal of Finance*, 48, 1803–1832.

WEIL, P. (1989): "The Equity Premium Puzzle and the Risk-free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–421.

——— (1992): "Equilibrium asset prices with undiversifiable labor income risk," *Journal of Economic Dynamics and Control*, 16, 769–790.

II.A Normalization

The optimization problem is the same for each agent. Thus, for notational convenience, we drop the agent specific superscript i in this section. Agents maximize utility (eq. (II.1) on page 44) with respect to the budget constraint (eq. (II.2) on page 45) and wealth accumulation (eq. (II.3) on page 45), i.e.

$$\max_{C_t,s_t,b_t} V_t(W_t, z_t) = \left[(C_t - \varsigma Y_t)^{1-\rho} + \beta \mathbb{E}_t \left[V_{t+1}(W_{t+1}, z_{t+1})^{1-\gamma}\right]^{\frac{1-\rho}{1-\gamma}}\right]^{\frac{1}{1-\rho}}, \tag{II.4}$$

s.t.

$$W_t + Y_{l,t} = Q_{s,t}s_t + Q_{b,t}b_t + C_t,$$

$$W_{t+1} = P_{s,t+1}s_t + P_{b,t+1}b_t.$$

To apply dynamic programming (Bellman (1957), Stokey and Lucas (1989)), we need our model to be stationary. To achieve this, we normalize all equations with the trending variable average income (Y_t). We denote the normalized variables in our model with lower case letters, i.e.

$$y_t = \tfrac{Y_t}{Y_t},\ v_t = \tfrac{V_t}{Y_t},\ v_{t+1} = \tfrac{V_{t+1}}{Y_{t+1}},\ w_t = \tfrac{W_t}{Y_t},\ q_{s,t} = \tfrac{Q_{s,t}}{Y_t},\ q_{b,t} = \tfrac{Q_{b,t}}{Y_t},\ c_t = \tfrac{C_t}{Y_t}.$$

The normalized payoffs to employees stock holders and bond holders are then

$$\begin{aligned}
y_{l,t} &= \frac{Y_{l,t}}{Y_t} = \frac{(1+\psi_t)\bar{l}\mathbb{E}_{t-1}[Y_t]}{Y_t} = \bar{l}(1+\psi_t)\frac{\mathbb{E}_{t-1}[g_t]}{g_t},\\
p_{b,t} &= \frac{P_{b,t}}{Y_t} = \frac{\mathbb{E}_{t-1}[Y_t]}{Y_t} = \frac{\mathbb{E}_{t-1}[g_t]}{g_t},\\
p_{s,t} &= \frac{P_{s,t}}{Y_t} = \frac{Y_t - (1-\bar{s})\mathbb{E}_{t-1}[Y_t]}{\bar{s}Y_t} = \frac{1-(1-\bar{s})\frac{\mathbb{E}_{t-1}[g_t]}{g_t}}{\bar{s}}.
\end{aligned}$$

We arrive at the normalized optimization problem by dividing (II.4) through Y_t. The normalized value function v_t has the additional factor g_{t+1} adjusting tomorrow's value to account for economic growth (see (II.5)). In the normalized version of our model, we choose to normalize prices and payoffs instead of asset holdings (see (II.6) and (II.7))

$$\max_{c_t,s_t,b_t} v_t(w_t, z_t) = \left[(c_t - \varsigma)^{1-\rho} + \beta \mathbb{E}_t \left[(g_{t+1}v_{t+1}(w_{t+1}, z_{t+1}))^{1-\gamma}\right]^{\frac{1-\rho}{1-\gamma}}\right]^{\frac{1}{1-\rho}}, \tag{II.5}$$

s.t.

$$w_t + y_t = q_{s,t}s_t + q_{b,t}b_t + c_t, \tag{II.6}$$

$$w_{t+1} = p_{s,t+1}s_t + p_{b,t+1}b_t. \tag{II.7}$$

II.B Equilibrium conditions

An analytic solution to (II.5) subject to (II.6) and (II.7) is unknown. To find a numeric solution, we solve the first order conditions using a nonlinear equation solver. In this section we derive the first order conditions starting from the Lagrangian.

II.B.1 Lagrangian

The Lagrangian for the normalized problem can be written as

$$\mathcal{L} = \left[(c_t - \varsigma)^{1-\rho} + \beta \mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{1-\gamma}\right]^{\frac{1-\rho}{1-\gamma}}\right]^{\frac{1}{1-\rho}} + \lambda \left[w_t + y_t - q_{s,t}s_t - q_{b,t}b_t - c_t\right]. \tag{II.8}$$

Note that we state the budget equation explicitly while substituting the wealth accumulation equation whenever necessary.

II.B.2 Derivatives

Differentiating (II.8) with respect to c_t and s_t yields

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial c_t} &= \frac{1}{1-\rho}\left[(c_t - \varsigma)^{1-\rho} + \beta \mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{1-\gamma}\right]^{\frac{1-\rho}{1-\gamma}}\right]^{\frac{1}{1-\rho}-1} (1-\rho)(c_t-\varsigma)^{-\rho} - \lambda = 0, \\
&= v_t^{\rho}(c_t - \varsigma)^{-\rho} - \lambda = 0,
\end{aligned} \tag{II.9}$$

$$\frac{\partial \mathcal{L}}{\partial s_t} = \frac{1}{1-\rho} v_t^{\rho} \beta \frac{1-\rho}{1-\gamma} \left(\mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{1-\gamma}\right]\right)^{\frac{1-\rho}{1-\gamma}-1} \mathbb{E}_t \left[g_{t+1}^{1-\gamma}(1-\gamma) v_{t+1}^{-\gamma} \frac{\partial v_{t+1}}{\partial c_{t+1}} \frac{\partial c_{t+1}}{\partial s_t}\right] - \lambda q_{s,t} = 0.$$

From (II.9) we see that $\frac{\partial v_{t+1}}{\partial c_{t+1}} = v_{t+1}^{\rho}(c_{t+1} - \varsigma)^{-\rho}$. The one-period lagged eq. (II.6) together with the wealth accumulation eq. (II.7) imply $\frac{\partial c_{t+1}}{\partial s_t} = p_{s,t+1}$. Then the derivative with respect to stock holdings can be written as

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial s_t} &= v_t^{\rho} \beta \left(\mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{1-\gamma}\right]\right)^{\frac{\gamma-\rho}{1-\gamma}} \mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{-\gamma} v_{t+1}^{\rho} (c_{t+1} - \varsigma)^{-\rho} p_{s,t+1}\right] - \lambda q_{s,t} = 0, \\
&= v_t^{\rho} \beta \left(\mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{1-\gamma}\right]\right)^{\frac{\gamma-\rho}{1-\gamma}} \mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{\rho-\gamma} (c_{t+1} - \varsigma)^{-\rho} p_{s,t+1}\right] - \lambda q_{s,t} = 0.
\end{aligned}$$

Differentiating with respect to bond holdings b_t yields

$$\frac{\partial \mathcal{L}}{\partial b_t} = v_t^{\rho} \beta \left(\mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{1-\gamma}\right]\right)^{\frac{\gamma-\rho}{1-\gamma}} \mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{\rho-\gamma} (c_{t+1} - \varsigma)^{-\rho} p_{b,t+1}\right] - \lambda q_{b,t} = 0.$$

II.B.3 Normalized first order conditions

Finally, we simplify the first order conditions normalizing by v_t^ρ. For this purpose define $\tilde{\lambda} = \frac{\lambda}{v_t^\rho}$. The first order conditions for consumption, stock holdings and bond holdings become

$$
\begin{aligned}
&(c_t - \varsigma)^{-\rho} - \tilde{\lambda} = 0, \\
&\beta \left(\mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{1-\gamma}\right]\right)^{\frac{\gamma-\rho}{1-\gamma}} \mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{\rho-\gamma} (c_{t+1} - \varsigma)^{-\rho} p_{s,t+1}\right] - \tilde{\lambda} q_{s,t} = 0, \\
&\beta \left(\mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{1-\gamma}\right]\right)^{\frac{\gamma-\rho}{1-\gamma}} \mathbb{E}_t \left[g_{t+1}^{1-\gamma} v_{t+1}^{\rho-\gamma} (c_{t+1} - \varsigma)^{-\rho} p_{b,t+1}\right] - \tilde{\lambda} q_{b,t} = 0.
\end{aligned}
$$

In the case of CRRA utility, $\rho = \gamma$. In this case, the first order conditions collapse to

$$
\begin{aligned}
&(c_t - \varsigma)^{-\gamma} - \tilde{\lambda} = 0, \\
&\beta \mathbb{E}_t \left[g_{t+1}^{1-\gamma} (c_{t+1} - \varsigma)^{-\gamma} p_{s,t+1}\right] - \tilde{\lambda} q_{s,t} = 0, \\
&\beta \mathbb{E}_t \left[g_{t+1}^{1-\gamma} (c_{t+1} - \varsigma)^{-\gamma} p_{b,t+1}\right] - \tilde{\lambda} q_{b,t} = 0.
\end{aligned}
$$

The full set of equilibrium conditions combines the above first order conditions with the market clearing conditions

$$
\sum_{i=1}^{n} s_t^i = \bar{s}, \quad \sum_{i=1}^{n} b_t^i = \bar{b}, \quad \forall t.
$$

II.C Computation

II.C.1 State space

The state of the economy is described by beginning of period normalized wealth w_t of all agents and the current shock z_t.

II.C.2 Smolyak

Solving for all agents but one requires to find a full solution of the model, since all agents have identical preferences. The remaining agent simply receives all residual quantities. Thus, the state space dimension is equal to the number of agents minus one. The main challenge in solving the model is the approximation of policy functions as they depend on the entire wealth vector w_t. We compute the case of six agents, thus the dimension of the continuous state space is five. Consider, for example, a course grid of five points per axis. In this case, the number of grid points already amounts to $5^5 = 3125$. The algorithm proposed in Smolyak (1963) has the advantage that the number of points grows polynomially rather than exponentially in the number of dimensions, providing a counterspell to the curse of dimensionality. Applications of the Smolyak approximation algorithm in the field of economics first appeared in Krueger and Kubler (2004). Recent applications are Krueger and Kubler (2006) in an overlapping-generations model and Malin, Krueger, and Kubler (2011) in a multi-country real business cycle model. To our knowledge, this is the first time, Smolyak approximation is applied to an infinite horizon competitive equilibrium model.

II.C.3 Implementation of the borrowing constraint

An intuitive implementation of the borrowing constraint $I_t \geq \underline{I}$ through Kuhn-Tucker conditions leads to kinks in the policy function. The Smolyak algorithm, however, requires the approximated function to be smooth. To avoid kinks, we compute an auxiliary problem.

Consider the normalized wealth space $[\underline{w}, \overline{w}]$ for each agent[8]. First, we compute next period's policy over the state space ignoring the borrowing constraint. To account for the constraint, we then overwrite the optimal policy with a constant outside of the bounds. This ensures the borrowing constraints lie exactly at the boundary of the wealth space.

The above procedure may be equivalently formulated by the constraint, $I_t \geq I(s_t, \underline{w}, w)$, where $I(.)$ is the investment policy and w refers to the wealth vector of all other agents. Economically, this constraint is cumbersome and has no straight forward interpretation. However, equivalence to the original problem can be achieved by choosing $\underline{I} = I(s_t^u, \underline{w}, w)$ in the converged policy, where $w = \frac{\underline{w}+\overline{w}}{2}$ and s_t^u the unemployed state. Assuming that the borrowing constraint is non-binding in any state, in which the agent is employed and that the agents investment policy does not depend on the others wealth vector, the auxiliary and original problem yield the same solution.

To enforce the desired constraint ($\underline{i}$) on normalized investment (i_t) reported in the paper, we employ a (costly) optimization technique to arrive at according grid lower bound ($\underline{w}$). We solve the model on average about seven times until we arrive at the desired lower bound.

II.C.4 Policy Iteration

We apply standard dynamic programming techniques as described in Judd (1998) to solve the model. As policies we choose to approximate the investment decision of each agent as a function of the state variables, specifying the exogenous shock and the wealth of each agent. Then we implement almost complete debt roll-over as the initial investment policy and initialize the value function for each agent as well as price policies for both assets, accordingly. Given these policies tomorrow, we solve for the optimal policies today at each Smolyak grid point and finally use the Smolyak algorithm to approximate the policies.

[8]This holds true for all agents except one. As described above, Smolyak requires a cubical state space. Thus, one agent's wealth space equalizes all others. In other words, the wealth space of one agent is the residual of all others, $[-(n-1)\overline{w}, -(n-1)\underline{w}]$. Economically this can be interpreted as one agent being capable to borrow almost infinite amounts.

We iterate backwards over time and repeat this procedure until investment policies, value functions and price policies are converged.

Finally, we simulate one million quarters and compute the empirical moments on individual consumption and asset returns reported in the paper.

Part III

Multivariate Markov Chain Approximations

Multivariate Markov Chain Approximations

Simon Scheuring*

May 23, 2012

Abstract

To solve equilibrium models numerically, it is necessary to discretize vector autoregressive processes (VAR) into a finite number of states. Univariate Markov chain approximations are well studied, however, few papers address the multivariate case. This paper presents three approaches to the problem: quadrature, moment matching and bin estimation.

Quadrature uses numerical integration schemes over the conditional distribution of the error terms. Moment matching replicates the first moments of the VAR. Bin estimation segments the data into bins and estimates the transition probabilities with maximum likelihood.

A comparative study in a standard asset pricing model shows that quadrature has difficulties when the model involves only few states, bin estimation fares better, while moment matching delivers the smallest errors. However, an experiment demonstrates the convincing results of moment matching to be double-edged. The introduction of a disaster state permits to alter model results despite matching all first and second moments.

*Department of Banking and Finance, University of Zurich, simon.scheuring@bf.uzh.ch. I am highly indebted to Felix Kubler for continuous support. Furthermore, I would like to express my greatest thanks to Benjamin Jonen with whom I initially started to work on Markov chain approximations. Without him, this paper would not exist. I am grateful to Johannes Brumm for helpful comments and I thank Edward Knotek II and Yikai Wang for letting me use their implementations of Tauchen (1986).

III.1 Introduction

There are two options to solve discrete-time dynamic models with a continuous state space numerically. First, apply a quadrature method over the exogenous processes. Second, approximate the exogenous process by a Markov chain with a finite number of states, commonly referred to as Markov chain approximation. The first option usually requires closed form solutions to the policy functions. However, as models in finance become more and more sophisticated, closed form policies are rarely available. Thus, it is of paramount interest to study the different possibilities to approach Markov chain approximations.

Let us first express the general idea more formally for the case of a bivariate vector autoregressive process (VAR). Assume the dynamics of continuous state processes y_t and z_t are given as

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{t,y} \\ \epsilon_{t,z} \end{pmatrix}, \qquad \epsilon_t \sim \mathcal{N}(0, \Sigma),$$

where α refers to the constant, θ to the persistence and $\mathcal{N}$ denotes the normal distribution with variance-covariance matrix Σ. The objective is to construct a shock (S) and transition (T) matrix, approximating y_t and z_t

$$\tilde{y} = \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \ldots \\ \tilde{y}_n \end{pmatrix}, \qquad \tilde{z} = \begin{pmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \ldots \\ \tilde{z}_n \end{pmatrix}, \qquad S = \begin{pmatrix} \tilde{y} & \tilde{z} \end{pmatrix}, \tag{III.1}$$

$$T = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix}, \tag{III.2}$$

where $\tilde{y}_i$ and $\tilde{z}_i$ refer to the values in state i and t_{ij} denotes the probability to jump from state i to state j. Despite its relevance for countless applications in economics and finance, this problem has so far been mainly studied in the univariate case.[1] To my knowledge this is the first study to discuss advantages and disadvantages of different approaches to the bivariate case.

The paper presents three alternatives to approach the problem. *Tauchen (1986)* and *Tauchen and Hussey (1991)* propose to imitate a numerical integration scheme over the normal distribution of the error terms. *Moment matching* provides an ad-hoc construction scheme for the shock and transition matrix to reproduce the first and second moments. Eventual overdetermination might be avoided by imposing symmetry conditions. *Bin estimation* estimates the shock and transition matrix from data in a two step procedure. First, I use a clustering method to separate the data into bins. Second, I estimate the entries of the transition matrix by maximum likelihood.

[1] For comparative analysis of Markov chain approximations to univariate autoregressive processes (AR) see Munk (1998), Burnside (2006), Floden (2008) and Kopecky and Suen (2010).

Following the presentation of the three approaches, I state a standard asset pricing model as a benchmark. It turns out that simple quadrature Tauchen (1986) has difficulties when only a small numbers of states is available. The problem becomes less severe with an increasing number of states, nevertheless, convergence to the true model result is not guaranteed. Although, theoretically convincing, Gaussian quadrature following Tauchen and Hussey (1991) lacks a stable implementation. Bin estimation fares somewhat better than the quadrature approaches. The best results are achieved using moment matching. Even for very small numbers of states the model result is very precise.

Subsequently, I show in an experiment that the very good results of moment matching might be doubled-edged. Usually, there are more entries in the shock and transition matrix than moments available. Therefore, constructing the moment matching Markov chain approximation is an overdetermined problem. Unless additional constraints are enforced, this may be exploited to generate varying model results, despite keeping the first and second moments constant.

III.2 Quadrature methods

Quadrature is the most commonly used approach to Markov chain approximations in applications. It was introduced by Tauchen (1986) with a simple quadrature rule and further extended in Tauchen and Hussey (1991) by using Gaussian quadrature. The intention of this section is to briefly describe the intuition and existing implementations of quadrature methods to Markov chain approximations.

III.2.1 Tauchen (1986)

Tauchen (1986) is based on the idea of a simple numerical integration rule. Figure III.1 illustrates the approach for the one-dimensional case. Let us assume without loss of generality that $\tilde{y}_1 < \tilde{y}_2 < \cdots < \tilde{y}_n$. The method requires the user to choose the size of the state space by providing the bandwidth m as a parameter. Then the boundaries of the state space are constructed as $\tilde{y}_1 = \mu_y - m\sigma_y$ and $\tilde{y}_n = \mu_y + m\sigma_y$, where μ_y denotes the unconditional mean and σ_y the unconditional standard deviation of the original process. The remaining values $y_2, \ldots, y_{n-1}$ are equispaced between the boundaries. Equipped with the discretization of the state space, integrating over the conditional normal distribution delivers the transition matrix.

For two or more dimensions, Tauchen (1986) simply proposes to do the same independently for each variable. Then assembling the separate discretizations gives the shock matrix and the transition matrix may be found accordingly by integrating over the conditional multidimensional normal distribution.

I was able to collect three different implementations of this method. They are all written in MATLAB®. The first is by Yikai Wang and Marcus Hagedorn (WH) of the University of Zurich. The second is by an unknown author, so simply referred to as Tauchen1986 (T) and the third accompanies Knotek and Terry (2008) (KT). The first two, WH and

Figure III.1: Markov chain approximation — Tauchen (1986)

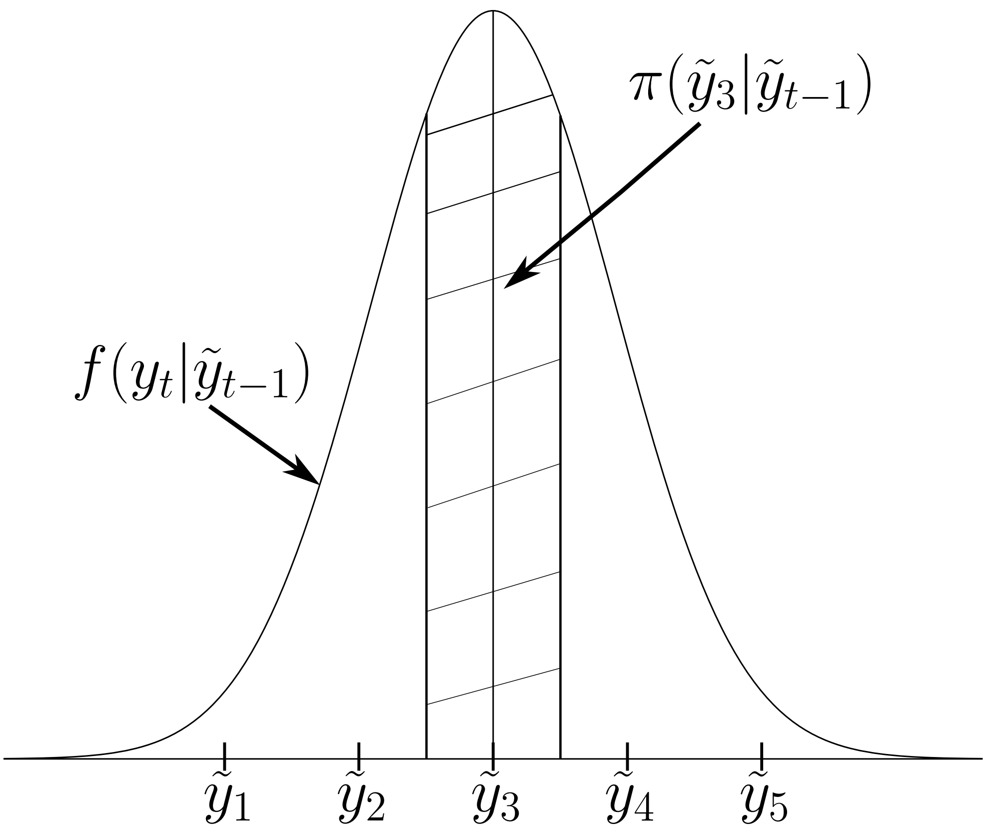


Illustration of the integration method for a univariate AR proposed in Tauchen (1986). $\tilde{y}_1$ to $\tilde{y}_5$ represent equally spaced entries in the shock vector. $f(y_t|\tilde{y}_{t-1})$ is the conditional distribution of y_t given the realization of last periods state y_{t-1}. The striped area, $\pi(\tilde{y}_3|\tilde{y}_{t-1})$, is the resulting transition probability to jump from state $\tilde{y}_{t-1}$ into state $\tilde{y}_3$.

T are pure implementations of Tauchen (1986). However, they employ a slightly different implementation of the construction of the shock matrix.[2] KT extends Tauchen, since it also allows for contemporaneous dependencies of the form

$$AX_t = \alpha + \theta X_{t-1} + \epsilon_t.$$

These dependencies do not allow for numerical integration. Therefore, KT switch to Monte Carlo instead. Therefore, convergence is slower than for the other two methods and the random draws cause a stochastic inaccuracy.

III.2.2 Tauchen and Hussey (1991)

One way to interpret Tauchen (1986) is too see it as a numerical integration rule. In this sense, it is very simplistic, since it takes equispaced points over the domain and attributes equal weights to each node. Tauchen and Hussey (1991) extends Tauchen (1986) consequently, by using the more sophisticated Gaussian quadrature. Essentially, Gaussian quadrature provides integrating points (y_k) and according weights (w_k), such that the integral of a function $g(y)$ against a density $w(y)$ is closely approximated

$$\int g(y)w(y)dy \approx \sum_{k=1}^{N} g(y_k)w_k.$$

Usually, in the one-dimensional case, abscissa y_k and weights y_k are chosen in such a way that the rule is exact for all polynomials with a degree smaller than $2N-1$. Theoretically, Gaussian quadrature is very convincing. Accordingly, in the one-dimensional case, this method is widely used in applications and has many advantages (Burnside (2006), Floden (2008)). However, in two or more dimensions Gaussian quadrature is quite tricky and still subject of ongoing research in numerics (Jäckel (2005) Taylor, Wingate, and Bos (2007)). As a consequence, to my knowledge, the only implementation capable of multidimensional Markov chain approximations, is the original one by George Tauchen in FORTRAN. It was until recently available on his website, but has now vanished. For the purpose of comparison, I wrote a MATLAB interface to FORTRAN.

III.3 Moment Matching

III.3.1 Idea

As almost any empiricist would confirm, error terms are usually not normally distributed in the data. Therefore, from a practical perspective, it seems like a futile effort to assume

[2] Sections III.B.1 and III.B.1 in the appendix show examples of the different shock matrices for the bivariate case with 9 states. WH space both variables independently on their respective grid and then construct the state space as all cross combinations. T makes the boundaries of the second variable contingent on the state of the first variable. Note that the approach by T will always lead to a wider range of the second variable.

normal distributions in the model and then approximate them with a quadrature method. This is particularly the case, when one faces computational limits and only few states are available. Therefore, in this section, I discuss an alternative, which abandons the distribution of the error terms and simply focuses on matching the first two moments, expectations, volatilities and correlations.

In the one-dimensional case of an autoregressive process (AR) of memory one, there is a construction approach by Rouwenhorst (1995). This approach has gained recent popularity through a comparative study by Kopecky and Suen (2010). This study shows that moment matching is very well suited for processes with high autocorrelation.

Unfortunately the two dimensional case suffers from over-determination for larger number of states. Therefore, I describe an ad-hoc approach to match first moments in two dimensions for 4 and 6 states. Although, this approach does not generalize, I believe it to be very useful when researchers face tight computational constraints.

III.3.2 Approach

In a first step, I construct a shock matrix that matches the first two unconditional moments: expectations, variances and correlation for an uniform unconditional distribution. Then I use a numerical solver to solve for a transition matrix that preserves the unconditional distribution and matches unconditional persistence as well as optional conditional moments.

Construction of the shock matrix

In the case of four states, I choose to impose symmetric deviations of the mean and construct the shock matrix as follows

$$y = \begin{pmatrix} \mu_y - \sigma_y \\ \mu_y - \sigma_y \\ \mu_y + \sigma_y \\ \mu_y + \sigma_y \end{pmatrix} \qquad z = \begin{pmatrix} \mu_z - \delta_1 \\ \mu_z - \delta_2 \\ \mu_z + \delta_2 \\ \mu_z + \delta_1 \end{pmatrix},$$

where μ_y and μ_z refer to the expectations and σ_y refers to the standard deviation of y. Expectations of both processes and the variance of y_t are precisely matched. Two parameters δ_1 and δ_2 remain to match the variance of z_t and the correlation between the two processes. This may be achieved with the formulas

$$\delta_2 = \frac{cov(y_t, z_t) + \sqrt{var(y_t)var(z_t) - cov(y_t, z_t)^2}}{\sigma_y},$$

$$\delta_1 = 2\frac{cov(y_t, z_t)}{\sigma_y} - \delta_2.$$

In the case of six states, I follow a similar approach for the discretization of the state space

$$\tilde{y} = \begin{pmatrix} \mu_y - \sqrt{\frac{3}{2}}\sigma_y \\ \mu_y - \sqrt{\frac{3}{2}}\sigma_y \\ \mu_y \\ \mu_y \\ \mu_y + \sqrt{\frac{3}{2}}\sigma_y \\ \mu_y + \sqrt{\frac{3}{2}}\sigma_y \end{pmatrix}, \qquad \tilde{z} = \begin{pmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \tilde{z}_3 \\ \tilde{z}_4 \\ \tilde{z}_5 \\ \tilde{z}_6 \end{pmatrix}.$$

The symmetry and the factor $\sqrt{\frac{3}{2}}$ jointly match mean and standard deviation of y. Then I use a solver to choose $\tilde{z}_1, ..., \tilde{z}_6$ such that they match the mean and volatility of z as well as the correlation between y and z. Obviously, the system is overdetermined, so it is possible to impose up to three additional constraints. One possibility would be too impose additional moments. However, for most datasets estimations of third or higher order moments tend to be highly imprecise. Alternatively, one could enforce either symmetries or economic conditions. I implemented two alternatives. Either fixing the outcome of specific states or enforcing symmetries[3].

Optimization of the transition matrix

Once the shock matrix is constructed, I find the accompanying transition matrix with a constrained optimization.

Constraints

First, the sum of each row has needs to be one to represent a probability

$$\sum_{j=1}^{n} t_{ij} = 1.$$

Second, I impose the transition matrix to be doubly-stochastic, e.g. the sum of each column is one

$$\sum_{i=1}^{n} t_{ij} = 1.$$

It is a well known result, that for any finite doubly-stochastic matrix, the unconditional distribution is uniform, e.g. each state has unconditional probability $\frac{1}{n}$. This has the advantage that the unconditional moments of the shock matrix remain unchanged.

Finally, I match the unconditional persistence, $\tilde{\theta} = \theta$. This requires to solve a linear system of equations, stated in appendix III.A.3 by equations (III.7) and (III.8).

[3] The symmetries are such that the lower half mirrors the top half around the mean (μ_z), e.g. $\tilde{z}_6 = \mu_z + (\mu_z - \tilde{z}_1)$, $\tilde{z}_5 = \mu_z + (\mu_z - \tilde{z}_2)$ and $\tilde{z}_4 = \mu_z + (\mu_z - \tilde{z}_3)$.

Objective

The constraints cover all truly necessary aspects. However, there still remain degrees of freedom. Basically, any objective on conditional moments might be considered. I choose to minimize the variance of conditional covariances. In most asset pricing models, once unconditional expectations are controlled for, conditional covariances are the main driver of asset prices. Therefore, to minimize the variance of conditional covariances leads to conservative transition matrices.

Optimization problem

The optimization problem then states:

$$\min_{t_{ij}} \mathbb{V}[cov_{t-1}(\tilde{y}_t, \tilde{z}_t)]$$

s.t.

$$\sum_{j=1}^{n} t_{ij} = 1$$
$$\sum_{i=1}^{n} t_{ij} = 1$$
$$\tilde{\theta} = \theta.$$

III.4 Bin estimation

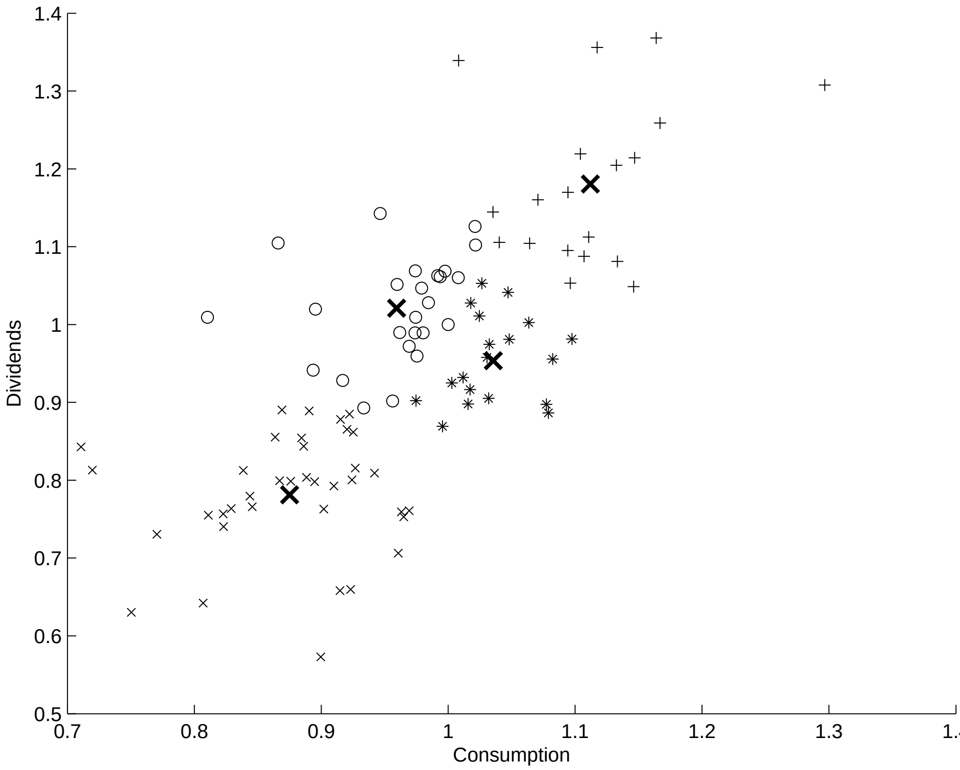
Moment matching avoids the estimation of the VAR by abandoning the distribution of error terms and directly matching empirical moments. In this section, I propose an approach to estimate the Markov chain directly from the empirical distribution. The procedure follows a two step approach. First, I cluster the data into segments. Then the shock matrix consists of the conditional expectation for each segment. In the second step, I estimate the transition matrix by maximum likelihood. Theoretically, the idea resembles Adda and Cooper (2003), except that I use the empirical distribution rather than making assumptions on the VAR.

III.4.1 Clustering — shock matrix

The first step is to segment the data into a number of clusters. Each cluster corresponds to one row (state) of the shock matrix. Any cluster method could be used for this task. I choose to use the in-built MATLAB routine *kmeans*. It "minimizes the sum, over all clusters, of the within-cluster sums of point-to-cluster-centroid [...] using squared Euclidian distances."[4]. This guarantees that nearby points are grouped together, which is the most important point from an economic perspective.

[4] MATLAB (2011) documentation.

Figure III.2: Bin estimation — four segments



An example of the clustering of 100 data points from a simulated time series over consumption and dividends following (III.3). The data is clustered into four separate segments, marked by circles, pluses, crosses and stars. The bold crosses (**x**) represent the center of their respective segment.

Figure III.2 shows an example of such a clustering. The crosses, stars, circles and pluses refer to groups of different data points. The bold **x** represent the center of each segment, used as the entry in the shock matrix. The data is simulated according to the process specified below in equation (III.3). The center of the groups are then the entries in the shock matrix. An example of such a shock matrix is shown in appendix III.B.4.

III.4.2 Maximum likelihood — transition matrix

Equipped with a shock matrix, the next step is to compute the transition matrix. Shalizi (2009) shows that the maximum likelihood estimation for transition probabilities comes down to simply counting the number of transitions from each group to the next. Consequently, the algorithm works as follows. Given the grouping into discrete states, the algorithm iterates through the empirical time series and counts every transition from one group to the next. Normalizing the counts by the number of data points delivers the transition matrix.

III.4.3 Bootstrap

When the data set is small relative to the number of bins used, the previous approach of counting transition often has the disadvantage that there are sometimes very few transitions from one group to another. The transition matrix might be imbalanced or even contain zeroes. In particular, zeroes might have undesirable economic implications, when agents adapt to the impossibility of certain events. This problem may be mitigated using bootstrap.

This is a three step procedure. First, I estimate the VAR process (see (III.1) on page 68) and obtain the coefficients $\hat{\alpha}, \hat{\theta}$ as well as the error terms $\hat{\epsilon} = \{\hat{\epsilon}_1, \hat{\epsilon}_2, ...\hat{\epsilon}_T\}$. Second, I consider a random variable u that has probability $\frac{1}{T}$ to take on each of the values of $\hat{\epsilon}$. Third, I construct a new time series by repeatedly drawing from the estimated error terms

$$y_t = \hat{\alpha} + \hat{\theta} y_{t-1} + u_t.$$

Then, I apply the previous to the new and prolonged time series.

III.5 Benchmark model

III.5.1 Motivation

The previous sections presented different approaches to Markov chain approximations. Theoretically, they all have advantages and disadvantages. Thus, it is very interesting to compare their performance in an application. I choose a standard asset pricing model. A representative agent faces two exogenous processes, consumption and dividends and prices two assets, stocks and bonds. The model has the advantage that it is relatively simple, so closed form solution for policy functions are readily available. Yet, despite its simplicity, the underlying dynamics are at the heart of many famous, well studied and successful models such as Mehra and Prescott (1985), Campbell and Cochrane (1999) or Bansal and Yaron (2004).

III.5.2 Model

The model is driven by a joint exogenous process for consumption (c_t) and dividends (d_t)

$$\begin{pmatrix} c_t \\ d_t \end{pmatrix} = \begin{pmatrix} \mu_c \\ \mu_d \end{pmatrix} + \begin{pmatrix} \theta_{cc} & \theta_{dc} \\ \theta_{cd} & \theta_{dd} \end{pmatrix} \begin{pmatrix} c_{t-1} \\ d_{t-1} \end{pmatrix} + \begin{pmatrix} \sigma_{cc} & \sigma_{dc} \\ \sigma_{cd} & \sigma_{dd} \end{pmatrix} \begin{pmatrix} \epsilon_{t,c} \\ \epsilon_{t,d} \end{pmatrix}, \quad \epsilon_t \sim \mathcal{N}(0,1). \tag{III.3}$$

One representative agent, prices two assets according to his utility function over the consumption stream

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t u(c_t)\right],$$

where the utility function has the standard form of constant relative risk aversion (CRRA) $u(c_t) = \frac{c^{1-\gamma}}{1-\gamma}$ with γ denoting risk aversion. Then the price of a one-period zero bond in the economy is easily derived as

$$q_t^B = \beta \frac{\mathbb{E}_t[u'(c_{t+1})]}{u'(c_t)}, \tag{III.4}$$

and similarly the price of a short-lived stock is given as[5]

$$q_t^D = \beta \frac{\mathbb{E}_t[u'(c_{t+1})d_{t+1}]}{u'(c_t)}. \tag{III.5}$$

Finally, I define the risk premium (RP) as the difference between the expected returns

$$RP = \mathbb{E}\left[\frac{\mathbb{E}_t[d_{t+1}]}{q_t^D} - \frac{1}{q_t^B}\right]. \tag{III.6}$$

III.5.3 Calibration

The purpose of this paper is to compare different approximation methods of exogenous variables, not to provide an explanation of the equity premium puzzle. I choose a calibration that provides substantial risk premia to make results easier to digest. Table III.1 shows the values. Arguably, the calibration is unreasonable with respect to volatilities and the correlation. A model providing an answer to the equity premium puzzle with a realistic calibration will necessarily exhibit much larger nonlinearities. Thus, in such a model approximation errors in the exogenous processes are likely to have an even larger impact on the model outcome.

[5] Implementing a long-lived stock would require to add the next period stock price to the dividends in the conditional expectation. Solving the model would then require to iterate over the state space. This complicates the solution of the model, however, does not add value to the purpose of comparing the different methods. Therefore, I make the simplifying assumption that the representative firm gets liquidated at the beginning of each period and is subsequently refounded.

Table III.1: Parameters

Parameter		Value
Discount factor	β	0.99
Risk aversion	γ	2
Persistence	θ_{cc}	0.30
	θ_{dc}	0.00
	θ_{cd}	0.15
	θ_{dd}	0.20
Mean	$\mathbb{E}[c_t]$	1.00
	$\mathbb{E}[d_t]$	1.00
Std. dev.	$\sigma[c_t]$	0.10
	$\sigma[d_t]$	0.15
Correlation	ρ	0.70

This table shows the calibration of the benchmark case. Parameters are chosen to generate substantial risk premia and do not attempt to be realistic. β discount factor; γ risk aversion; θ persistence; c consumption; d dividends; $\mathbb{E}$ expectation; σ standard deviation; ρ correlation.

III.6 Comparison

The asset pricing model provides the testing ground for the comparison. I obtain a precise solution to the problem with direct integration. Then, I compare the model results for different implementations of Tauchen (1986) for many states. Finally, I compare moments and model results for all methods for smaller number of states.

III.6.1 Integration

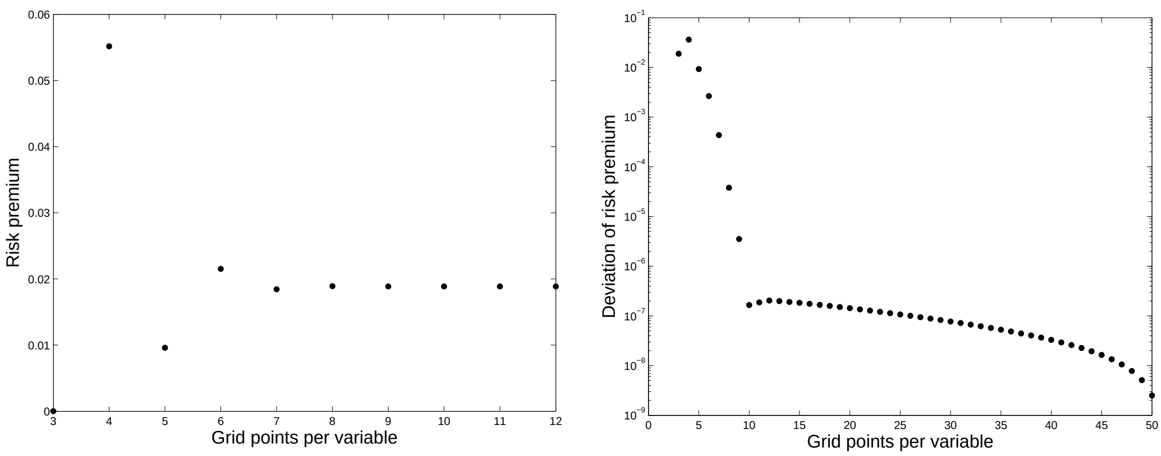
Solving the model by integration serves to obtain the true model solution of the continuous model as a benchmark value. This makes the exercise of Markov chain approximations redundant for this model. However, solving by direct integration is only possible, since the model is kept very simplistic and closed form solutions for the policies are available.

Computing the state-contingent prices requires a one-dimensional integration over the conditional distribution of consumption to solve for the bond (III.4) and a two-dimensional integration over consumption and dividends for the stock (III.5). Then, integrating the conditional prices over the unconditional distributions of income and dividends delivers the expected stock and bond returns. Finally, the risk premium (III.6) is easily derived as the difference between the two.

Since computational resources are not an issue here, I choose a basic and very stable integration method: pick boundaries as unconditional mean plus five times the standard

Figure III.3: Integration: risk premium

(a) Risk premium **(b) Deviation**



The left subfigure shows the risk premium obtained by direct integration as the number of grid points increases. The right figure shows the absolute deviation, where the reference value is the model result for 51 grid points.

deviation[6] and choose an equispaced grid, then I simply sample at each grid point and some them up weighted with the density of the normally distributed innovations. Figure III.3a shows how the risk premium converges over an increase in the number of integration points. Figure III.3b shows on a log scale, how the deviation converges. The deviation converges very quickly to 10^{-7} as the number of integration points per variable reaches 10. Thus, the result is shown to be very accurate and may be used as a benchmark value for the comparison of the different Markov chain approximations.

III.6.2 Comparison — many states

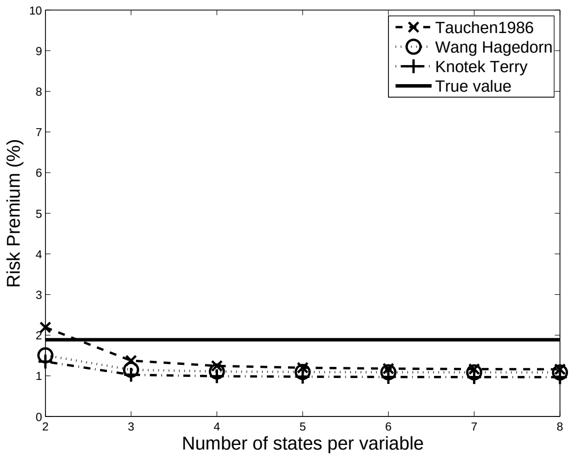
Only the three implementations of Tauchen (1986) are capable of handling large numbers of states. The implementation of Tauchen and Hussey (1991) does not converge for more than four states per variable, moment matching is only implemented for four and six total states and the clustering part of bin estimation is unstable for large numbers of states. Therefore, I compare the three implementations of Tauchen (1986) separately for many states and then proceed with a comparison of all methods for few states.

Figure III.4 shows the resulting risk premium compared to the number of states per variable. The different subfigures represent different bandwidths, e.g. sizes of the state space. The solid black line represents the true value as solved by integration. The three dotted lines represent different implementations of Tauchen (1986).

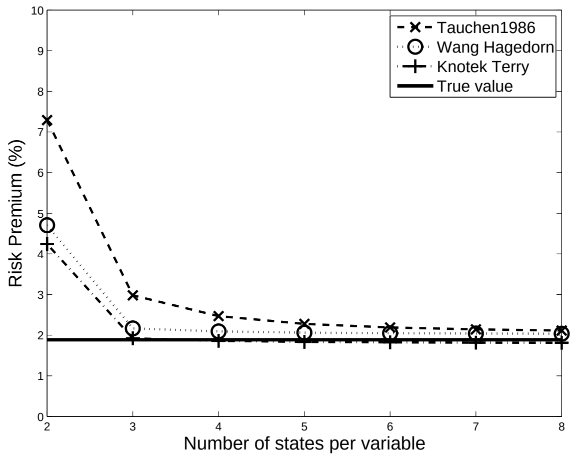
[6] As the exogenous variables are normally distributed, five times the standard distribution implies that the probability mass lying outside the grid is less than 5×10^{-7} for each variable.

Figure III.4: Comparison of different implementations of Tauchen (1986)

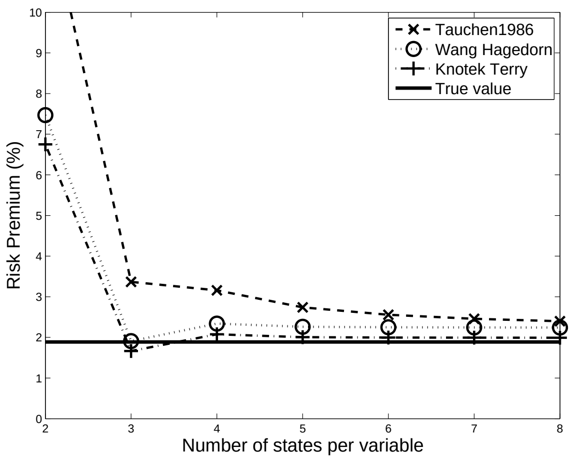
(a) Bandwidth $= 1$



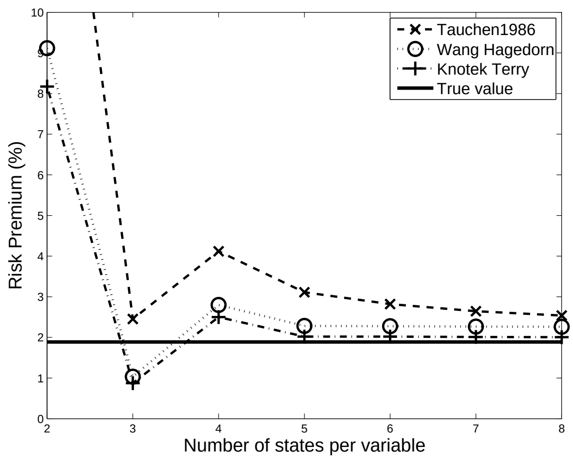
(b) Bandwidth $= 2$



(c) Bandwidth $= 3$



(d) Bandwidth $= 4$



These figures compare the model results of three implementations of Tauchen (1986) with the true value. The subfigures report different values for the bandwidth, e.g. the size of the state space as a multiple of the standard deviation.

Subfigure III.4a displays a bandwidth of one, i.e. the boundary points are mean plus standard deviation. This leads to risk premia, which are always too low, since the state space approximation is too narrow. The only exception is for two states per variable. Then a bandwidth of 1 actually provides the best results.

A bandwidth of 2, displayed in Subfigure III.4b is the default parametrization of Knotek Terry (KT). Indeed, it provides almost perfect results for KT. For the other two methods the risk premium is too high and slowly tends toward the true value. Surprisingly, for larger bandwidths, Subfigures III.4c and III.4d, the results do not really improve. For smaller numbers of states the risk premium is far off, for larger values convergence is slow.

These figures demonstrate that it is extremely dangerous to use only 2 states per variable. For two states, the otherwise too narrow bandwidth of 1 is even the best choice, however, risk premia might still be anything between 1.3% and 2.2%. These results suggest that for reasonably close results a bandwidth of 2 and at least 5 states per variable, so 25 states in total should be used.

Unfortunately, when solving sophisticated equilibrium models, researchers often do not have enough computing time to use so many discrete states. The next subsection compares different methods for fewer states.

III.6.3 Comparison — few states

Moments

Table III.2 shows the moments for various approximation methods. The first column indicates the broader group of the approximation method. The next column indicates the precise method used, followed by the total number of states.

For the three Tauchen methods, I used a bandwidth of $\sqrt{3/2}$, since there are 9 states in total, e.g. three states per variable. As described in section III.3.2 on moment matching this bandwidth matches the unconditional standard deviation exactly, if the transition matrix were doubly stochastic. Consequently, the standard deviations are close to the true value. All three methods have the same result for the persistence, which is reasonable close, yet not exact. However, correlations vary due to the different construction approaches of the shock matrix. The approach to specify a conditional secondary column of Tauchen1986 appears to be more suited to match the correlation. Surprisingly, the theoretically superior method of Tauchen and Hussey (1991) has difficulties to match the moments. Independently of the number of states used, the results are relatively far off. Most likely, this reflects problems with the implementation of multidimensional Gaussian quadrature. By construction, moment matching delivers the exact results. Finally, bin estimation is designed to use actual data rather than approximate an assumed VAR. Thus, I simulated 10000 data points prior to applying bin estimation. The resulting moments slightly underestimate standard deviations and overstate the correlation. I believe this phenomenon is systematic due to the clustering. Taking the center of each cluster attributes less weight to outliers. However, as more states are used, the moments get much closer to the actual values. For 9 states even the persistence is reasonably approximated.

Table III.2: Comparison of moments for small numbers of states

Method		θ_{cc}	θ_{dc}	θ_{cd}	θ_{dd}	$\sigma[c_t]$	$\sigma[d_t]$	ρ
True value		0.30	0.00	0.15	0.20	0.10	0.15	0.70
Tauchen	T9	0.25	−0.00	0.12	0.16	0.09	0.14	0.71
	WH9	0.25	0.00	0.12	0.16	0.09	0.14	0.59
	KT9	0.25	−0.00	0.12	0.16	0.09	0.13	0.60
Tauchen Hussey	TH4	0.03	−0.00	−0.06	0.11	0.10	0.18	0.57
	TH9	0.03	−0.00	−0.06	0.11	0.10	0.18	0.57
	TH16	0.03	0.00	−0.06	0.11	0.10	0.18	0.57
Moment	MM4	0.30	−0.00	0.15	0.20	0.10	0.15	0.70
	MM6	0.30	−0.00	0.15	0.20	0.10	0.15	0.70
Bins	BE4	0.12	0.05	0.16	0.08	0.08	0.13	0.83
	BE6	0.19	0.01	0.17	−0.02	0.09	0.13	0.81
	BE9	0.24	0.05	0.16	0.19	0.09	0.14	0.77

This table compares the first moments for different Markov chain approximations. The first line repeats the assumed values of the VAR from Table III.1. The first column indicates the broader group of approximation methods. The second column indicates the precise method, followed by the total number of states, so T9 stands for Tauchen with nine states. The following columns report the different moments. The underlying shock matrices may be found in appendix III.B.
The used abbreviations are: T, Tauchen 1986; WH, Wang Hagedorn; KT, Knotek Terry; TH, Tauchen Hussey; MM, Moment matching; BE, Bin estimation; θ, persistence; σ, standard deviation; ρ, correlation.

Table III.3: Comparison of model results for small numbers of states

Method		r_B	%dev.	r_D	%dev.	RP	%dev.
Integration		1.0031	0%	1.0220	0.0%	0.0189	0.0%
Tauchen	T9	1.0034	0.03%	1.0214	−0.06%	0.0180	−4.64%
	WH9	1.0034	0.03%	1.0182	−0.37%	0.0148	−21.37%
	KT9	1.0043	0.12%	1.0176	−0.43%	0.0132	−29.78%
Tauchen Hussey	TH4	1.0089	0.58%	1.0304	0.82%	0.0214	13.47%
	TH9	1.0089	0.57%	1.0312	0.90%	0.0224	18.48%
	TH16	1.0088	0.57%	1.0312	0.90%	0.0224	18.73%
Moment	MM4	1.0018	−0.14%	1.0208	−0.12%	0.0190	0.90%
	MM6	1.0018	−0.13%	1.0207	−0.13%	0.0189	0.16%
Bins	BE4	1.0018	−0.13%	1.0207	−0.13%	0.0173	−8.31%
	BE6	1.0051	0.19%	1.0247	0.26%	0.0196	3.94%
	BE9	1.0033	0.02%	1.0214	−0.05%	0.0181	−3.84%

This table compares the model results and deviations of the true value for different Markov chain approximations. The first line states the true solution of the model as obtained by direct integration. The first column indicates the broader group of approximation methods. The second column indicates the precise method, followed by the total number of states. The following columns report the results of the different approximation methods and the according percent deviation of the true value. The underlying shock matrices may be found in appendix III.B.
The used abbreviations are: T, Tauchen 1986; WH, Wang Hagedorn; KT, Knotek Terry; TH, Tauchen Hussey; MM, Moment matching; BE, Bin estimation; r_B, riskfree rate; r_D, stock return; RP, risk premium.

Model results

Table III.3 shows the model results for the riskfree rate r_B, the stock return r_D and the risk premium for the different Markov chain approximations, discussed in this paper. The first row states the precise results as obtained by direct integration in section III.6.1. The following rows relate to the different approximation methods.

Among the implementations of Tauchen (1986) the deviations of the risk premium vary from 5% to 30%. This is a surprising and important result, since they all employ the same method and the methods differ only in the construction of the shock matrix. Tauchen and Hussey (1991) has difficulties of a similar magnitude with deviations between 13% and 19%. Surprisingly, the results get worse for nine and sixteen total states, despite unchanged moments in Table III.2. This indicates once again difficulties with multivariate Gaussian quadrature.

Moment matching leads to excellent results. Despite the very small number of states used, the deviations from the true value of the risk premium are below 1%. This is very promising, however, may partly be an artifact of the simplicity of the model. It is unclear, whether this result would hold in a model with stronger nonlinearities. Finally, bin estimation has decent results with risk premia between 9% and 4%. It is somewhat worrying that the results alternate from positive to negative deviations as the number of states increases. The cause is probably the stochastic nature of the approach.

III.7 Moment matching — disaster state

In the previous comparative study, moment matching had very impressive results, coming very close to the true value with very few states. Nevertheless, there remains the conceptual problem that moment matching ignores the distribution of the error terms. In this section, I demonstrate that this is dangerous and may be exploited.

III.7.1 Experiment

The shock matrix construction for six states has three degrees of freedom. In the previous section, I circumvented this problem, by imposing symmetry conditions. In this section, I perform an experiment and demonstrate that the model results can vary dramatically, although all moments are precisely matched.

Let us denote the payoff in the worst growth state (disaster), as d_d. Then, I construct the discretization of the shock matrix as:

$$\tilde{c} = \begin{pmatrix} \mu_c - \sqrt{\frac{3}{2}}\sigma_c \\ \mu_c - \sqrt{\frac{3}{2}}\sigma_c \\ \mu_c \\ \mu_c \\ \mu_c + \sqrt{\frac{3}{2}}\sigma_c \\ \mu_c + \sqrt{\frac{3}{2}}\sigma_c \end{pmatrix} \qquad \tilde{d} = \begin{pmatrix} d_d \\ d_d \\ \tilde{d}_3 \\ \tilde{d}_4 \\ \tilde{d}_5 \\ \mu_d + (\mu_d - d_d) \end{pmatrix},$$

The first two entries of the dividend vector $\tilde{d}$ correspond to the lowest growth state and I thus refer to them as the disaster state. As described in section III.3.2 there are three degrees of freedom in the dividend matrix. Therefore, I impose a symmetry in the last state and use the remaining three parameters, $\tilde{d}_3, \ldots, \tilde{d}_5$ to match the mean and standard deviation of dividends and the correlation between dividends and consumption. This construction guarantees that independently of d_d, the first two moments are matched. The transition matrix follows by the constrained optimization described in section III.3.2.

III.7.2 Moments

Table III.4 shows the moments over varying disaster states (d_d). The first row repeats the true values from the parameterization in Table III.1. The next panel shows the moments for moment matching with 4 and 6 states. For 6 states, symmetries as described in section III.3.2 are imposed. The third panel displays the moments over different disaster states ranging from 0.8 to 0.95. For disaster state values between 0.825 and 0.925 all the moments

Table III.4: Disaster state: moments

Method		d_d	θ_{cc}	θ_{dc}	θ_{cd}	θ_{dd}	$\sigma[c_t]$	$\sigma[d_t]$	ρ
True value			0.30	0.00	0.15	0.20	0.10	0.15	0.70
Original	MM4		0.30	−0.00	0.15	0.20	0.10	0.15	0.70
	MM6		0.30	−0.00	0.15	0.20	0.10	0.15	0.70
Disaster	MM6	0.800	0.30	−0.00	0.15	0.20	0.10	0.17	0.63
	MM6	0.825	0.30	0.00	0.15	0.20	0.10	0.15	0.70
	MM6	0.850	0.30	−0.00	0.15	0.20	0.10	0.15	0.70
	MM6	0.875	0.30	−0.00	0.15	0.20	0.10	0.15	0.70
	MM6	0.900	0.30	−0.00	0.15	0.20	0.10	0.15	0.70
	MM6	0.925	0.30	0.00	0.15	0.20	0.10	0.15	0.70
	MM6	0.950	0.30	0.00	0.15	0.20	0.10	0.18	0.59

This table demonstrates that the moments remain unchanged, although one state of the shock matrix changes. d_d refers to the dividend payoff in the disaster state. The first row displays the true values. The second panel repeats the moments of Table III.2. The third panel reports the moments for different disaster states.

are the same. Only for the two values 0.8 and 0.95 the algorithm has difficulties to match the correlations.

III.7.3 Results

Table III.5 shows the model results as a different stock payoff in the worst consumption state is enforced. The risk free rate does not change. However, stock returns vary dramatically. As the payoff of the stock in the worst consumption state increases, the stock becomes less "risky" to the representative agent and the stock return diminishes. Accordingly the risk premium decreases as well. Remarkable are the large deviations in the risk premium, ranging from −10% to 10%, although the moments are entirely unchanged over that range. Figure III.5 displays the percent deviation of the risk premium over the disaster state. It shows that the percent deviation decreases monotonically and almost linearly as the stock payoff in the worst consumption state increases.

III.8 Conclusion

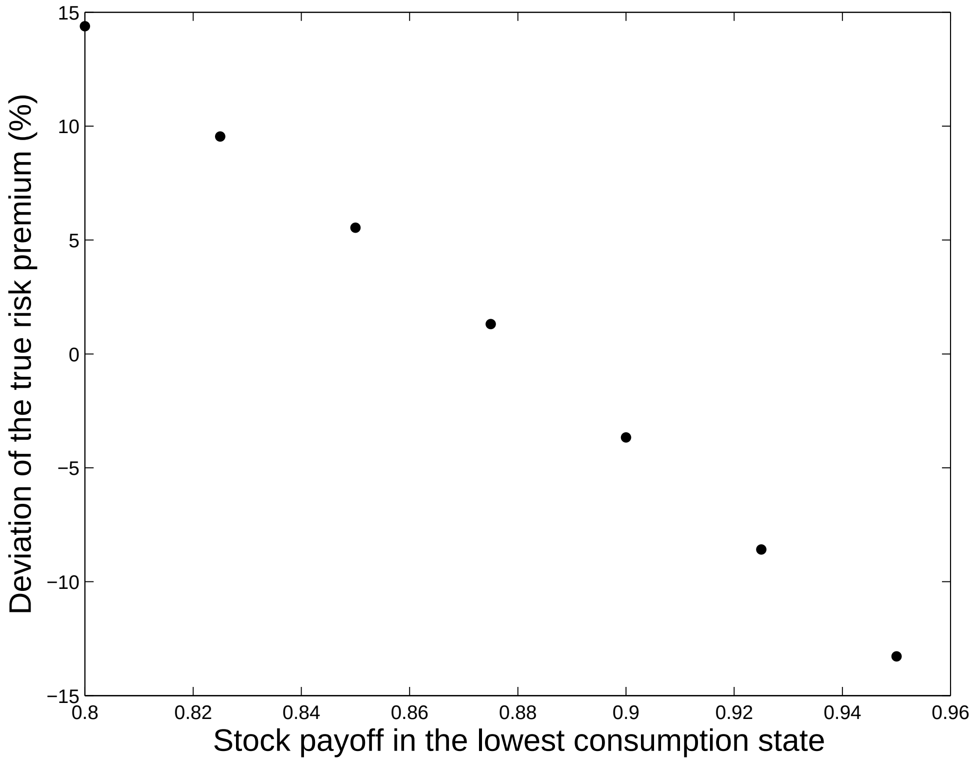
This paper discussed three methods for Markov chain approximations of two dimensional VAR processes. Quadrature, moment matching and bin estimation are compared in a standard financial economic model. Quadrature following Tauchen (1986) has difficulties for small numbers of states and is highly sensitive to the arbitrary choice of a multiple to determine the range of the state space. Tauchen and Hussey (1991) is theoretically promising, however the only available implementation appears to have stability issues. Moment

Table III.5: Disaster state: model results

Method		d_d	r_B	$\%dev.$	r_D	$\%dev.$	RP	$\%dev.$
Integration			1.0031	0%	1.0220	0.0%	0.0189	0.0%
Original	MM4		1.0018	−0.14%	1.0208	−0.12%	0.0190	0.90%
	MM6		1.0018	−0.13%	1.0207	−0.13%	0.0189	0.16%
Disaster	MM6	0.800	1.0016	−0.15%	1.0232	0.12%	0.0216	14.38%
	MM6	0.825	1.0017	−0.14%	1.0224	0.04%	0.0207	9.54%
	MM6	0.850	1.0017	−0.15%	1.0216	−0.04%	0.0199	5.54%
	MM6	0.875	1.0017	−0.15%	1.0208	−0.12%	0.0191	1.31%
	MM6	0.900	1.0017	−0.15%	1.0198	−0.21%	0.0182	−3.67%
	MM6	0.925	1.0017	−0.15%	1.0189	−0.30%	0.0172	−8.59%
	MM6	0.950	1.0016	−0.15%	1.0180	−0.39%	0.0164	−13.28%

This table shows the change in the model result as I vary the stock payoff in the state with lowest consumption d_d.

Figure III.5: Disaster state: percent deviation of the risk premium



This figure shows how moment matching may be abused to generate different risk premia, while keeping the first two moments constant. The abscissa reports different stock payoffs in the worst consumption state. The ordinate shows the percent deviation of the resulting risk premia.

matching provides excellent model results, however, is overdetermined for a larger number of states. When one does not employ additional conditions, this overdetermination may be exploited to generate varying model results, while matching all first and second order moments. It might be a worthwhile direction of future research to develop a sensible construction approach for more states. Finally, bin estimation as the most intuitive approach delivers reasonably good results. However, not as great as moment matching, and when datasets are small, imbalances of the transition matrix might become an issue.

The central message of this paper is that researchers need to be very careful when applying Markov chain approximations. Something as trivial as choosing a different implementation of Tauchen (1986) may have a significant impact on model results. Furthermore, detailed documentation is necessary, in particular, only reporting the resulting moments is not sufficient.

Bibliography

ADDA, J., AND R. COOPER (2003): *Dynamic Economics: Quantitative Methods and Applications.* MIT Press, Cambridge, MA.

BANSAL, R., AND A. YARON (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59(4), 1481–1509.

BURNSIDE, C. (2006): "Discrete State-Space Methods for the Study of Dynamic Economies," WP.

CAMPBELL, J. Y., AND J. H. COCHRANE (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *The Journal of Political Economy*, 107(2), 205–251.

FLODEN, M. (2008): "A note on the accuracy of Markov-chain approximations to highly persistent AR(1) processes," *Economics Letters*, 99(3), 516–520.

JÄCKEL, P. (2005): "A note on multivariate Gauss-Hermite quadrature," http://www.jaeckel.org/.

KNOTEK, E. S., AND S. TERRY (2008): "Markov-chain approximations of vector autoregressions: application of general multivariate-normal integration techniques," Research Working Paper RWP 08-02, Federal Reserve Bank of Kansas City.

KOPECKY, K. A., AND R. M. SUEN (2010): "Finite state Markov-chain approximations to highly persistent processes," *Review of Economic Dynamics*, 13, 701–714.

MATLAB (2011): *version 7.12.0 (R2011a).* The MathWorks Inc., Natick, Massachusetts.

MEHRA, R., AND E. C. PRESCOTT (1985): "The equity premium: A puzzle," *Journal of Monetary Economics*, 15(2), 145 – 161.

MUNK, C. (1998): "The Markov Chain Approximation Approach for Numerical Solution of Stochastic Control Problems: Experiences from Merton's Problem," WP.

ROUWENHORST, K. (1995): *Frontiers of Business Cycle Research*chap. Asset Pricing Implications of Equilibrium Business Cycle Models, pp. 294–330. Princeton University Press.

Shalizi, C. (2009): "Note: Maximum Likelihood Estimation for Markov Chains," 34-462.

Tauchen, G. (1986): "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20(2), 177 – 181.

Tauchen, G., and R. Hussey (1991): "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59(2), 371–396.

Taylor, M., B. Wingate, and L. Bos (2007): "A cardinal function algorithm for compution multivariate quadrature points," *SIAM Journal of Numerical Analysis*, 45, 193–205.

III.A Closed form formulas for moment matching

This section provides formulas to compute moments for a given shock matrix S (III.1) and transition matrix T (III.2). The results are mainly used in the moment matching section, but also serve, when computing the results for Tables III.2 and III.4, which report moments of different Markov chain approximations.

Note, in the paper I use $\sim$ to mark discrete approximations and to clearly distinguish them from their continuous counterparts. In this appendix everything is discrete and therefore I omit the $\sim$ for notational ease.

III.A.1 Stationary distribution

The probabilities of the stationary distribution λ of the transition matrix T may be found by either simply computing T^{∞} or by solving the left eigenvector of T, which corresponds to the Eigenvalue 1, since

$$\lambda = \lambda T.$$

III.A.2 First two moments

Conditional Mean

$$\bar{y}_i = \sum_{j=1}^{n} t_{ij} y_j$$

$$\bar{z}_i = \sum_{j=1}^{n} t_{ij} z_j$$

Unconditional Mean

$$\bar{y} = \sum_{i=1}^{n} \lambda_i y_i$$

$$\bar{z} = \sum_{i=1}^{n} \lambda_i z_i$$

Unconditional Variance

$$\sigma_y^2 = \sum_{i=1}^{n} \lambda_i (y_i - \bar{y})^2$$

$$\sigma_z^2 = \sum_{i=1}^{n} \lambda_i (z_i - \bar{z})^2$$

Unconditional Covariance

$$\sigma_{yz}^2 = \sum_{i=1}^{n} \lambda_i (y_i - \bar{y})(z_i - \bar{z})$$

Conditional Covariance

$$E(y_t z_t | s_{t-1} = s_i) = \sum_{j=1}^{N} t_{ij} y_j z_j$$
$$cov(y_t, z_t | s_{t-1} = s_i) = E(y_t z_t | s_{t-1} = s_i) - E(y_t | s_{t-1} = s_i) E(z_t | s_{t-1} = s_i)$$

Correlation

$$\rho_{yz} = \frac{\sigma_{yz}}{\sigma_y \sigma_z}$$

III.A.3 Persistence

Unfortunately there is no generic closed form expression for the persistence. However, it may be found by solving a linear system of equations.

First compute the following moments:

$$\begin{aligned} E(y_{t+1} y_t) &= E(E(y_{t+1} y_t | y_t = y_i)) \\ &= \sum_{i=1}^{n} \lambda_i y_i E(y_{t+1} | y_t = y_i) \\ &= \sum_{i=1}^{n} \lambda_i y_i \sum_{j=1}^{N} t_{ij} y_j \\ &= \sum_{i=1}^{n} \lambda_i y_i \bar{y}_i \end{aligned}$$

$$\begin{aligned} E(y_{t+1} z_t) &= E(E(y_{t+1} z_t | z_t = z_i)) \\ &= \sum_{i=1}^{n} \lambda_i z_i E(y_{t+1} | z_t = z_i) \\ &= \sum_{i=1}^{n} \lambda_i z_i \sum_{j=1}^{N} t_{ij} y_j \\ &= \sum_{i=1}^{n} \lambda_i z_i \bar{y}_i \end{aligned}$$

Second, solve the following equation system to get the persistence coefficients

$$\frac{1}{n}\begin{pmatrix} 1 & E(y_{t-1}) & E(z_{t-1}) \\ E(y_{t-1}) & E(y_{t-1}^2) & E(y_{t-1}z_{t-1}) \\ E(z_{t-1}) & E(z_{t-1}y_{t-1}) & E(z_{t-1}^2) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \theta_{11} \\ \theta_{12} \end{pmatrix} = \frac{1}{n}\begin{pmatrix} E(y_t) \\ E(y_t y_{t-1}) \\ E(y_t z_{t-1}) \end{pmatrix} \tag{III.7}$$

$$\frac{1}{n}\begin{pmatrix} 1 & E(y_{t-1}) & E(z_{t-1}) \\ E(y_{t-1}) & E(y_{t-1}^2) & E(y_{t-1}z_{t-1}) \\ E(z_{t-1}) & E(z_{t-1}y_{t-1}) & E(z_{t-1}^2) \end{pmatrix} \begin{pmatrix} \alpha_2 \\ \theta_{21} \\ \theta_{22} \end{pmatrix} = \frac{1}{n}\begin{pmatrix} E(z_t) \\ E(z_t y_{t-1}) \\ E(z_t z_{t-1}) \end{pmatrix} \tag{III.8}$$

III.B Selected shock matrices

These shock matrices accompany Tables III.2 and III.3. The specific layout of the shock matrix has a crucial impact on model results. Therefore, this section lists all Shock matrices of the different methods used in the comparative study.

III.B.1 Tauchen

Tauchen 1986 — 9 states — bandwidth $\sqrt{3/2}$ (T9)

$$\begin{pmatrix} 0.87 & 0.73 \\ 1.00 & 0.87 \\ 1.13 & 1.00 \\ 0.87 & 0.87 \\ 1.00 & 1.00 \\ 1.13 & 1.13 \\ 0.87 & 1.00 \\ 1.00 & 1.13 \\ 1.13 & 1.27 \end{pmatrix}$$

Wang Hagedorn — 9 states — bandwidth $\sqrt{3/2}$ (WH9)

$$\begin{pmatrix} 0.87 & 0.81 \\ 1.00 & 0.81 \\ 1.13 & 0.81 \\ 0.87 & 1.00 \\ 1.00 & 1.00 \\ 1.13 & 1.00 \\ 0.87 & 1.19 \\ 1.00 & 1.19 \\ 1.13 & 1.19 \end{pmatrix}$$

Knotek Terry — 9 states — bandwidth $\sqrt{3/2}$ (KT9)

$$\begin{pmatrix} 0.88 & 0.82 \\ 0.88 & 1.00 \\ 0.88 & 1.18 \\ 1.00 & 0.82 \\ 1.00 & 1.00 \\ 1.00 & 1.18 \\ 1.12 & 0.82 \\ 1.12 & 1.00 \\ 1.12 & 1.18 \end{pmatrix}$$

III.B.2 Tauchen Hussey

Tauchen Hussey — 4 states (TH4)

$$\begin{pmatrix} 0.90 & 0.74 \\ 1.10 & 0.95 \\ 0.90 & 1.04 \\ 1.10 & 1.25 \end{pmatrix}$$

Tauchen Hussey — 9 states (TH9)

$$\begin{pmatrix} 0.83 & 0.56 \\ 1.00 & 0.74 \\ 1.17 & 0.92 \\ 0.83 & 0.82 \\ 1.00 & 1.00 \\ 1.17 & 1.18 \\ 0.83 & 1.08 \\ 1.00 & 1.26 \\ 1.17 & 1.44 \end{pmatrix}$$

Tauchen Hussey — 16 states (TH16) is omitted, since it takes up a lot of space and is structurally the same as TH9.

III.B.3 Moment matching

Moment matching — 4 states (MM4)

$$\begin{pmatrix} 0.90 & 0.79 \\ 0.90 & 1.00 \\ 1.10 & 1.00 \\ 1.10 & 1.21 \end{pmatrix}$$

Moment matching — 6 states (MM6)

$$\begin{pmatrix} 0.88 & 0.75 \\ 0.88 & 0.99 \\ 1.00 & 0.93 \\ 1.00 & 1.07 \\ 1.12 & 1.01 \\ 1.12 & 1.25 \end{pmatrix}$$

III.B.4 Bin estimation

Bin estimation — 4 states (BE4)

$$\begin{pmatrix} 0.89 & 0.82 \\ 0.97 & 1.05 \\ 1.03 & 0.95 \\ 1.12 & 1.17 \end{pmatrix}$$

Bin estimation — 6 states (BE6)

$$\begin{pmatrix} 0.85 & 0.78 \\ 0.94 & 1.00 \\ 0.97 & 0.86 \\ 1.03 & 1.14 \\ 1.06 & 1.00 \\ 1.15 & 1.21 \end{pmatrix}$$

Bin estimation — 9 states (BE9)

$$\begin{pmatrix} 0.84 & 0.74 \\ 0.89 & 0.92 \\ 0.94 & 1.06 \\ 0.96 & 0.84 \\ 1.01 & 1.00 \\ 1.04 & 1.16 \\ 1.07 & 0.94 \\ 1.12 & 1.09 \\ 1.17 & 1.27 \end{pmatrix}$$

Simon Scheuring – Curriculum Vitae

Date of Birth: 25.07.1984
Place of Birth: Filderstadt, Germany

Educational Background

10/2008 – 09/2012	**University of Zurich, Switzerland** Doctoral Studies in Economics and Business Administration Specialization in Banking and Finance
10/2007 – 09/2008	**University of Oxford, United Kingdom** MSc. Mathematical Modelling and Scientific Computing
09/2005 – 07/2007	**ENSAE, Paris, France** Diplôme de Statisticien-Economiste Double Diploma program with the University of Mannheim
10/2002 – 07/2007	**University of Mannheim, Germany** Diplom-Volkswirt
08/1995 – 03/2002	**Carl-Bosch Gymnasium Ludwigshafen, Germany** Abitur

Work Experience

09/2008 – 08/2012	**University of Zurich, Switzerland** Research Assistant at the Department of Banking and Finance
06/2007 – 09/2008	**WestLB AG, London, United Kingdom** Internship in Credit Structuring
04/2002 – 12/2002	**Bundeswehr, schweres Pionierbataillon, Speyer, Germany** Military service